HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Nguyễn Thành Nam

NGHIÊN CỨU VÀ PHÁT TRIỂN HỆ THỐNG NHẬN DẠNG NGÔN NGỮ KÝ HIỆU NGƯỜI CÂM ĐIẾC ỨNG DỤNG TRONG THƯ VIỆN SỐ

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT
(Theo định hướng ứng dụng)

HÀ NỘI – 2025

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Nguyễn Thành Nam

NGHIÊN CỨU VÀ PHÁT TRIỂN HỆ THỐNG NHẬN DẠNG NGÔN NGỮ KÝ HIỆU NGƯỜI CÂM ĐIẾC ỨNG DỤNG TRONG THƯ VIỆN

CHUYÊN NGÀNH : HỆ THỐNG THÔNG TIN

MÃ SỐ : 8.48.01.04 (Hệ thống thông tin)

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC: TS. PHAN THỊ HÀ

HÀ NỘI – 2025

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



Nguyễn Thành Nam

NGHIÊN CỨU VÀ PHÁT TRIỂN HỆ THỐNG NHẬN DẠNG NGÔN NGỮ KÝ HIỆU NGƯỜI CÂM ĐIẾC ỨNG DỤNG TRONG THƯ VIỆN

CHUYÊN NGÀNH : HỆ THỐNG THÔNG TIN

MÃ SỐ : 8.48.01.04 (Hệ thống thông tin)

ĐỀ ÁN TỐT NGHIỆP THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

NGƯỜI HƯỚNG DẪN KHOA HỌC: TS. PHAN THỊ HÀ

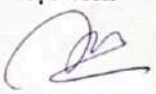
HÀ NỘI – 2025

LỜI CAM ĐOAN

Tôi xin khẳng định rằng đây là kết quả nghiên cứu độc lập của tôi dưới sự hướng dẫn tận tình của TS. Phan Thị Hà.

Kết quả và các số liệu nghiên cứu trong đề án tốt nghiệp của tôi là trung thực, đảm bảo tính khoa học, khách quan và chưa từng được mọi người công bố trong bất kỳ công trình khoa học nào khác mà học viên được biết. Toàn bộ các tài liệu tham khảo, thông tin trích dẫn trong đề án tốt nghiệp được ghi nguồn gốc rõ ràng.

Học viên



Nguyễn Thành Nam

LỜI CẢM ƠN

Lời đầu tiên tôi xin bày tỏ lời cảm ơn chân thành nhất tới TS. Phan Thị Hà đã tận tình chỉ bảo và tạo mọi điều kiện để tôi có thể hoàn thành đề án tốt nghiệp của mình một cách tốt nhất.

Tôi xin gửi lời cảm ơn tới các Thầy Cô giáo khoa Sau Đại học – Học viện Công nghệ Bưu chính Viễn thông đã có những hướng dẫn, giúp đỡ tôi trong suốt thời gian tôi tham gia học tập tại trường, đặc biệt là trong giai đoạn tôi làm đề án tốt nghiệp.

Tôi cũng xin gửi lời cảm ơn tới toàn thể Ban Lãnh đạo Trung tâm giải pháp Giáo dục điện tử đã giúp đỡ và tạo điều kiện cho tôi trong quá trình thu thập thông tin và tìm kiếm tài liệu để tôi có thể hoàn thành đề án tốt nghiệp của mình một cách hoàn chỉnh nhất.

Trong quá trình thực hiện, do còn vài hạn chế về lý luận, kinh nghiệm, thời gian nghiên cứu nên đề án tốt nghiệp không thể tránh khỏi những sai sót. Tôi rất mong nhận được những ý kiến đóng góp của các Thầy, Cô giáo và các bạn để đề án tốt nghiệp của tôi được hoàn thiện hơn.

Tôi xin chân thành cảm ơn.

Học viên

Nguyễn Thành Nam

MỤC LỤC

LỜI CAM ĐOAN.....1
LỜI CẢM ƠN.....2
DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT.....6
DANH MỤC CÁC BẢNG..... 7
DANH MỤC CÁC HÌNH.....8
MỞ ĐẦU.....1
1. Lý do chọn đề tài..... 1
2. Tổng quan về vấn đề nghiên cứu.....3
3. Mục đích nghiên cứu..... 4
4. Đối tượng và phạm vi nghiên cứu..... 4
5. Phương pháp nghiên cứu:..... 5
6. Kết cấu của đề án.....6
CHƯƠNG 1:
CƠ SỞ LÝ LUẬN..... 7
1.1. Các khái niệm liên quan đến nhận dạng ngôn ngữ ký hiệu..... 7
1.1.1. Khái niệm ngôn ngữ ký hiệu..... 7
1.1.2. Khái niệm nhận dạng ngôn ngữ ký hiệu..... 8
1.1.3. Nhận dạng ngôn ngữ ký hiệu trong hệ thống thư viện số..... 8
1.1.4. Khái niệm về học máy (Machine Learning)..... 9
1.1.5. Khái niệm về học sâu (Deep Learning)..... 9
1.1.5. Khái niệm về thị giác máy tính (Computer Vision)..... 10
1.1.6. Khái niệm về nhận dạng cử chỉ và ngôn ngữ ký hiệu..... 11
1.2. Tổng quan nghiên cứu trong nước và ngoài nước..... 11
1.2.1. Các nghiên cứu trên thế giới..... 11
1.2.2. Các nghiên cứu tại Việt Nam..... 12
1.2.3. Nhận xét và khoảng trống nghiên cứu..... 12
1.3. Định hướng nghiên cứu của đề án..... 13
CHƯƠNG 2:
PHÂN TÍCH, THIẾT KẾ VÀ PHÁT TRIỂN HỆ THỐNG NHẬN DẠNG NGÔN NGỮ KÝ HIỆU ỨNG DỤNG TRONG THƯ VIỆN SỐ..... 15
2.1. Phân tích yêu cầu hệ thống..... 15
2.1.1. Phân tích đối tượng sử dụng hệ thống.....15
2.1.2. Phân tích bối cảnh và hệ thống thư viện số hiện tại.....17
2.1.3. Xác định yêu cầu chức năng của hệ thống.....19

2.1.4. Xác định yêu cầu phi chức năng..... 21
2.2. Phân tích bài toán nhận dạng ngôn ngữ ký hiệu.....23
2.2.1. Mô tả bài toán..... 23
2.2.1. Đặc thù của ngôn ngữ ký hiệu Việt Nam trong bài toán nhận dạng.... 24
2.3. Thiết kế tổng thể hệ thống..... 25
2.3.1. Nguyên tắc thiết kế hệ thống.....25
2.3.2. Kiến trúc tổng thể của hệ thống..... 26
2.3.3. Thiết kế luồng xử lý nghiệp vụ..... 26
2.3.4. Thiết kế giao tiếp giữa các thành phần..... 27
2.4. Xây dựng và xử lý bộ dữ liệu ngôn ngữ ký hiệu Việt Nam.....28
2.4.1. Xác định phạm vi và nội dung bộ dữ liệu..... 28
2.4.2. Quy trình thu thập dữ liệu..... 31
2.4.3. Tiền xử lý dữ liệu..... 33
2.4.2. Gán nhãn và tổ chức dữ liệu..... 34
2.5. Huấn luyện và tối ưu mô hình học máy.....36
2.5.1. Lựa chọn và triển khai thuật toán.....36
2.5.2. Phương pháp 1: Học máy truyền thống (Random Forest)..... 38
2.5.3. Phương pháp 2: Mô hình CNN-LSTM..... 39
2.5.4. Phương pháp 3: Mô hình LSTM trên dữ liệu điểm mốc.....41
2.5.5. Đánh giá và so sánh các mô hình.....43
Lý do lựa chọn mô hình cuối cùng..... 45
2.6. Triển khai và tích hợp hệ thống..... 45
2.6.1. Môi trường triển khai..... 45
2.6.2. Triển khai dịch vụ nhận dạng ngôn ngữ ký hiệu.....47
2.6.3. Tích hợp với hệ thống thư viện số..... 47
2.6.1. Tối ưu hiệu năng và bảo mật.....49
CHƯƠNG 3:
ĐÁNH GIÁ VÀ THỬ NGHIỆM HỆ THỐNG NHẬN DẠNG NGÔN NGỮ KÝ HIỆU VIỆT NAM..... 51
3.1. Mục tiêu và phương pháp đánh giá hệ thống..... 51
3.2. Thiết lập kịch bản thử nghiệm và đánh giá..... 52
3.2.1. Đối tượng và phạm vi thử nghiệm..... 52
3.2.2. Các kịch bản thử nghiệm chính.....53
3.2.3. Điều kiện và quy trình thử nghiệm..... 53
3.2.4. Tiêu chí đánh giá trong thử nghiệm..... 54
3.3. Các chỉ số và phương pháp đánh giá hiệu quả hệ thống..... 54

3.3.1. Độ chính xác nhận dạng (Accuracy).....54
3.3.2. Độ chính xác theo lớp (Precision, Recall và F1-score).....55
3.3.3. Thời gian phản hồi của hệ thống.....55
3.3.4. Tỷ lệ nhận dạng sai và nhu cầu thực hiện lại ký hiệu..... 56
3.3.5. Đánh giá trải nghiệm người dùng..... 56
3.4. Kết quả thử nghiệm và phân tích đánh giá..... 56
3.4.1. Kết quả đánh giá độ chính xác của các mô hình.....57
3.4.2. Phân tích kết quả nhận dạng theo nhóm ký hiệu.....57
3.4.3. Đánh giá thời gian phản hồi của hệ thống.....58
3.4.4. Đánh giá khả năng sử dụng và trải nghiệm người dùng..... 58
3.4.5. Nhận xét tổng hợp kết quả thử nghiệm.....59
3.5. Đánh giá tổng thể và kết luận chương..... 59
KẾT LUẬN.....61
Các kết quả đạt được của đề án:..... 61
Nhận xét và đánh giá..... 61
Đề xuất và khuyến nghị..... 62
Hướng nghiên cứu tiếp theo..... 62
DANH MỤC CÁC TÀI LIỆU THAM KHẢO..... 63

DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

Viết tắt	Tiếng Anh	Giải thích
PTIT	Posts and Telecommunications Institute of Technology	Học viện Công nghệ Bưu chính Viễn thông
WCAG	Web Content Accessibility Guidelines	Bộ hướng dẫn khả năng tiếp cận nội dung web
VSL	Vietnamese Sign Language	Ngôn ngữ ký hiệu Việt Nam
AI	Artificial Intelligence	Trí tuệ nhân tạo
ML	Machine Learning	Học máy
DL	Deep Learning	Học sâu
CNN	Convolutional Neural Network	Mạng nơ-ron tích chập
RNN	Recurrent Neural Network	Mạng nơ-ron hồi tiếp
LSTM	Long Short-Term Memory	Mạng bộ nhớ dài – ngắn hạn
API	Application Programming Interface	Giao diện lập trình ứng dụng
REST	Representational State Transfer	Kiến trúc giao tiếp dịch vụ web
UI	User Interface	Giao diện người dùng
UX	User Experience	Trải nghiệm người dùng
FPS	Frames Per Second	Số khung hình trên giây

DANH MỤC CÁC BẢNG

Bảng 2.1. Thống kê lượng dữ liệu xây dựng bộ dữ liệu mẫu 33
Bảng 2.2. So sánh độ chính xác giữa các mô hình 43
Bảng 2.3. Đánh giá tổng hợp các mô hình theo mục tiêu đề án 44
Bảng 2.4. Cấu hình phần cứng đáp ứng huấn luyện 46
Bảng 2.5. Các yêu cầu phần mềm đáp ứng triển khai 46
Bảng 3.1. So sánh độ chính xác của các mô hình nhận dạng 55
Bảng 3.2. Độ chính xác theo nhóm dữ liệu 55
Bảng 3.3. Thời gian phản hồi trung bình của hệ thống 56

DANH MỤC CÁC HÌNH

Hình 1.1. Tổng quan mô hình thị giác máy tính 10
Hình 1.2. Thủ ngữ - Ngôn ngữ ký hiệu tay theo chữ cái latin 11
Hình 2.1. Hình ảnh trang chủ hệ thống thư viện số do VNPT-IT phát triển 15
Hình 2.2. Hình ảnh mô tả chức năng của người dùng với hệ thống thư viện số 18
Hình 2.3. Giao diện chức năng phòng học đặc biệt 48
Hình 2.4. Giao diện chi tiết phòng học đặc biệt tích hợp camera và thân thiện người dùng 49

MỞ ĐẦU

1. Lý do chọn đề tài

Trong bối cảnh chuyển đổi số và xây dựng xã hội thông tin ngày càng phát triển mạnh mẽ, việc đảm bảo quyền tiếp cận thông tin bình đẳng cho mọi người dân trở thành một yêu cầu cấp thiết. Cộng đồng người khiếm thính tại Việt Nam với khoảng 1,5 triệu người đang đối mặt với nhiều rào cản trong việc tiếp cận thông tin và dịch vụ công, đặc biệt là các dịch vụ thư viện số và tra cứu tài liệu trực tuyến. Ngôn ngữ ký hiệu Việt Nam là ngôn ngữ tự nhiên và phương tiện giao tiếp chính của cộng đồng người khiếm thính, tuy nhiên hầu hết các hệ thống thông tin số hiện nay đều được thiết kế dựa trên ngôn ngữ nói và chữ viết, tạo ra khoảng cách lớn giữa người khiếm thính và nguồn tri thức xã hội.

Hệ thống thư viện số thuvienso.digilib.vn của Trung tâm Giải pháp giáo dục – Công ty công nghệ thông tin – VNPT-IT hiện đang quản lý hàng trăm nghìn tài liệu số phục vụ nhu cầu học tập và nghiên cứu của cộng đồng. Tuy nhiên, với giao diện và phương thức tìm kiếm chủ yếu dựa vào văn bản tiếng Việt, hệ thống này chưa thực sự thân thiện và dễ tiếp cận đối với người khiếm thính, những người thường gặp khó khăn trong việc đọc hiểu tiếng Việt do đây không phải là ngôn ngữ tự nhiên đầu tiên của họ. Điều này không chỉ hạn chế quyền tiếp cận tri thức của một bộ phận công dân mà còn là một thách thức lớn trong việc thực hiện cam kết về xây dựng xã hội bao trùm và bình đẳng.

Yếu tố công nghệ trở thành nhân tố quyết định trong việc xóa bỏ rào cản và tạo điều kiện bình đẳng cho người khuyết tật. Công nghệ nhận dạng ngôn ngữ ký hiệu dựa trên trí tuệ nhân tạo và thị giác máy tính đã có những bước phát triển đáng kể trên thế giới, đặc biệt là đối với các ngôn ngữ ký hiệu phổ biến như ASL (Mỹ) hay BSL (Anh). Tuy nhiên, đối với ngôn ngữ ký hiệu Việt Nam, lĩnh vực nghiên cứu này vẫn còn rất hạn chế do thiếu nguồn dữ liệu chất lượng cao và các nghiên cứu ứng dụng thực tế.

Tài sản quý giá nhất của một xã hội phát triển chính là sự bình đẳng trong tiếp cận tri thức và thông tin. Để đảm bảo quyền lợi chính đáng của người khiếm

thính, cần phải tận dụng và phối hợp nhiều nguồn lực, trong đó công nghệ thông tin và trí tuệ nhân tạo là công cụ không thể thiếu. Thực tế cho thấy, các quốc gia nào đầu tư vào công nghệ hỗ trợ người khuyết tật và biết cách ứng dụng hiệu quả thì sẽ xây dựng được xã hội bao trùm và phát triển bền vững. Bỏ qua yếu tố công nghệ, chỉ dựa vào các phương thức truyền thống như phiên dịch viên ngôn ngữ ký hiệu hay phụ đề thủ công là chưa đủ để đáp ứng nhu cầu ngày càng cao của cộng đồng người khiếm thính trong thời đại số.

Việc xây dựng hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam không chỉ giúp người khiếm thính tiếp cận thông tin một cách tự nhiên và thuận tiện hơn mà còn góp phần quan trọng trong việc thúc đẩy hòa nhập xã hội, nâng cao chất lượng cuộc sống và tạo cơ hội học tập, phát triển bình đẳng cho cộng đồng này. Đây có ý nghĩa thực tiễn rất lớn không chỉ đối với hệ thống thư viện số mà còn có thể mở rộng cho các dịch vụ công khác như y tế, giáo dục, hành chính công.

Hệ thống thư viện số thuvienso.digilib.vn, sau gần 10 năm hoạt động, đã bước vào giai đoạn nâng cấp và mở rộng dịch vụ. Trước đây, các nhà phát triển chủ yếu tập trung vào việc số hóa tài liệu, quản lý kho dữ liệu và tối ưu hóa công cụ tìm kiếm văn bản. Tuy nhiên, trong giai đoạn hiện nay, việc nâng cao khả năng tiếp cận và trải nghiệm người dùng đa dạng được chú trọng hơn, xem đây là yếu tố then chốt để thư viện số thực sự phục vụ toàn thể cộng đồng. Do đó, việc tích hợp công nghệ nhận dạng ngôn ngữ ký hiệu Việt Nam vào hệ thống tìm kiếm đang được đặt ra như một mục tiêu quan trọng trong lộ trình phát triển.

Tuy nhiên, thực tế cho thấy việc triển khai công nghệ này vẫn gặp nhiều thách thức. Thứ nhất, chưa có bộ dữ liệu ngôn ngữ ký hiệu Việt Nam chuyên biệt cho lĩnh vực thư viện và tra cứu thông tin. Thứ hai, các nghiên cứu hiện có chủ yếu dừng lại ở giai đoạn thử nghiệm trong phòng thí nghiệm, chưa được tích hợp vào hệ thống thực tế. Thứ ba, thiếu sự tham gia của cộng đồng người khiếm thính trong quá trình thiết kế và phát triển, dẫn đến các giải pháp có thể không đáp ứng nhu cầu thực tiễn. Vì vậy, cần xây dựng một giải pháp tích hợp toàn diện, từ thu thập dữ

liệu, phát triển mô hình nhận dạng đến triển khai hệ thống và đánh giá với người dùng thực tế.

Do vậy, tôi chọn đề tài **"Nghiên cứu và phát triển hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam cho người khiếm thính ứng dụng trong thư viện số"** làm đề tài luận văn thạc sĩ của mình, đồng thời cũng mong muốn đóng góp một giải pháp công nghệ thiết thực giúp cộng đồng người khiếm thính có thể tiếp cận tri thức dễ dàng hơn, đồng thời cung cấp cho các nhà phát triển hệ thống thông tin một tài liệu tham khảo cả về lý thuyết lẫn phương pháp triển khai công nghệ hỗ trợ người khuyết tật.

2. Tổng quan về vấn đề nghiên cứu

Nhận dạng ngôn ngữ ký hiệu là một lĩnh vực nghiên cứu quan trọng của thị giác máy tính, hướng tới việc tự động chuyển các cử chỉ tay và chuyển động cơ thể thành văn bản hoặc tín hiệu có ý nghĩa. Trên thế giới, lĩnh vực này đã phát triển đa dạng, từ các phương pháp trích xuất đặc trưng truyền thống đến các mô hình học sâu hiện đại như mạng tích chập, mạng hồi quy và kiến trúc Transformer. Ở Việt Nam, những năm gần đây đã xuất hiện các nghiên cứu bước đầu về nhận dạng ngôn ngữ ký hiệu Việt Nam, tập trung chủ yếu vào việc xây dựng bộ dữ liệu và nhận dạng ký hiệu đơn trong điều kiện ánh sáng và bối cảnh được kiểm soát, cho kết quả tương đối khả quan.

Tuy vậy, vẫn tồn tại nhiều hạn chế chưa được giải quyết triệt để. Phần lớn các nghiên cứu hiện nay chỉ dừng lại ở từ vựng cơ bản, chưa đáp ứng nhu cầu nhận dạng các cụm từ phục vụ tra cứu thông tin trong thư viện số. Bên cạnh đó, việc đánh giá mô hình chủ yếu được thực hiện trong môi trường phòng thí nghiệm, khiến hiệu quả suy giảm khi áp dụng vào điều kiện sử dụng thực tế với nhiều yếu tố gây nhiễu như ánh sáng không ổn định, nền phức tạp hay góc quay thay đổi. Ngoài ra, các nghiên cứu tích hợp công nghệ nhận dạng ký hiệu vào hệ thống thư viện số đang vận hành gần như chưa được đề cập.

Xuất phát từ những khoảng trống trên, đề tài này tập trung vào ba hướng chính: xây dựng bộ dữ liệu ngôn ngữ ký hiệu Việt Nam chuyên biệt cho hoạt động

tìm kiếm tài liệu; phát triển mô hình nhận dạng có khả năng hoạt động ổn định trong môi trường thực tế; và triển khai giao diện tương tác tích hợp trực tiếp vào nền tảng thuvienso.digilib.vn nhằm cải thiện khả năng tiếp cận thông tin cho cộng đồng người khiếm thính.

3. Mục đích nghiên cứu

Mục đích nghiên cứu của đề án là nghiên cứu, thiết kế và xây dựng hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam tích hợp vào thư viện số thuvienso.digilib.vn, nhằm nâng cao khả năng tiếp cận và sử dụng tài nguyên số cho cộng đồng người khiếm thính.

Để đạt được mục đích nghiên cứu, đề án tập trung thực hiện các mục tiêu cụ thể sau:

– Hệ thống hóa cơ sở lý luận và các nghiên cứu liên quan đến nhận dạng ngôn ngữ ký hiệu, đặc biệt là ngôn ngữ ký hiệu Việt Nam trong bối cảnh ứng dụng công nghệ thị giác máy tính và học máy;

– Phân tích thực trạng khả năng tiếp cận thông tin của người khiếm thính trên các hệ thống thư viện số, đồng thời đánh giá hiện trạng kỹ thuật của thư viện số thuvienso.digilib.vn;

– Đề xuất, thiết kế và triển khai giải pháp nhận dạng ngôn ngữ ký hiệu Việt Nam phục vụ chức năng tìm kiếm tài liệu, tích hợp vào hệ thống thư viện số thuvienso.digilib.vn, đáp ứng yêu cầu sử dụng trong môi trường thực tế.

4. Đối tượng và phạm vi nghiên cứu

- ***Đối tượng nghiên cứu của đề án*** là ngôn ngữ ký hiệu Việt Nam dùng trong hoạt động tìm kiếm tài liệu trên thư viện số; Hệ thống thuvienso.digilib.vn và các thành phần kỹ thuật liên quan và Nhóm người dùng khiếm thính có nhu cầu sử dụng thư viện số.
- ***Phạm vi nghiên cứu của đề án:***

- Phạm vi về không gian: triển khai trong phạm vi Việt Nam, tập trung tại Hà Nội, TP.HCM, Đà Nẵng.

- Phạm vi về thời gian: Dữ liệu được phân tích chủ yếu trong khoảng thời gian 6 tháng (02/2025–05/2025).

5. Phương pháp nghiên cứu:

- ***Phương pháp thu thập dữ liệu***: Số liệu được thu thập từ các nguồn sau:

Thu thập dữ liệu thứ cấp: Là các tài liệu kỹ thuật, tài liệu thiết kế kiến trúc, mô tả chức năng, API và dữ liệu vận hành của nền tảng thư viện số *thuvienso.digilib.vn*. Các tài liệu này giúp phân tích hiện trạng hệ thống, khả năng tích hợp và mở rộng thêm module nhận dạng ngôn ngữ ký hiệu. Việc thu thập dữ liệu nội bộ giúp quá trình nghiên cứu được thực hiện nhanh chóng, thuận lợi và bảo đảm tính chính xác..

Dữ liệu thứ cấp bên ngoài: Nghiên cứu, tham khảo các công trình khoa học, bài báo, luận văn trong và ngoài nước liên quan đến nhận dạng ngôn ngữ ký hiệu, thị giác máy tính, học sâu và các mô hình như CNN[], LSTM[], Transformer[]. Đồng thời tham khảo các tiêu chuẩn, chính sách và nghiên cứu về khả năng tiếp cận thông tin cho người khiếm thính trong môi trường số, làm cơ sở so sánh và đề xuất giải pháp phù hợp.

Thu thập dữ liệu sơ cấp: Tiến hành xây dựng bộ dữ liệu ngôn ngữ ký hiệu Việt Nam phục vụ cho tác vụ tìm kiếm tài liệu trong thư viện số. Dữ liệu được thu thập từ người khiếm thính thông qua việc ghi hình các ký hiệu tay và cụm từ thường dùng trong tra cứu thư viện. Quá trình thu thập được thực hiện trong nhiều điều kiện khác nhau về ánh sáng, nền cảnh và góc quay nhằm phản ánh sát môi trường sử dụng thực tế.

Cơ sở xây dựng bộ dữ liệu và kịch bản thu thập dựa trên các nghiên cứu trước về nhận dạng ngôn ngữ ký hiệu và yêu cầu chức năng của hệ thống thư viện số. Dữ liệu sau khi thu thập được tiền xử lý, gán nhãn và chuẩn hóa để phục vụ huấn luyện mô hình học máy.

Nội dung dữ liệu thu thập gồm:

- Thông tin về ký hiệu đơn lẻ và các cụm từ thường dùng trong tìm kiếm thư viện.

- Chuỗi tọa độ các điểm mốc bàn tay theo từng khung hình.

- Nhãn dữ liệu tương ứng với từng ký hiệu và cụm từ.

- ***Phương pháp xử lý dữ liệu:***

Dữ liệu thu thập được xử lý thông qua các bước tiền xử lý như lọc nhiễu, chuẩn hóa tọa độ, cân bằng dữ liệu và chia tập huấn luyện – kiểm tra. Các kết quả nhận dạng được thống kê và đánh giá chủ yếu theo phương pháp định lượng kết hợp định tính nhằm phản ánh chính xác hiệu quả của mô hình.

- ***Phương pháp xử lý thông tin:***

Sử dụng phương pháp tổng hợp, phân tích và so sánh. Các số liệu về độ chính xác, hiệu năng mô hình và kết quả nhận dạng được xử lý và phân tích bằng các công cụ hỗ trợ (Python, Excel) để đánh giá mức độ đáp ứng yêu cầu của hệ thống và đề xuất hướng cải thiện.

- ***Phương pháp đối chiếu, so sánh:***

So sánh kết quả nghiên cứu của đề án với các công trình liên quan trong và ngoài nước về nhận dạng ngôn ngữ ký hiệu; đồng thời đối chiếu kết quả mô hình giữa các giai đoạn huấn luyện và thử nghiệm để đánh giá tính ổn định và khả năng ứng dụng trong môi trường thư viện số thực tế.

6. Kết cấu của đề án

Ngoài phần mở đầu, kết luận và tài liệu tham khảo, đề án được trình bày theo 3 chương gồm:

Chương 1: Cơ sở lý luận.

Chương 2: Phân tích, thiết kế và phát triển hệ thống.

Chương 3: Triển khai thử nghiệm và đánh giá hệ thống.

CHƯƠNG 1:
CƠ SỞ LÝ LUẬN

Người dùng, đặc biệt là cộng đồng người khiếm thính, giữ vai trò trung tâm quyết định hiệu quả và giá trị thực tiễn của các hệ thống thư viện số. Khả năng tiếp cận thông tin một cách chủ động và thuận tiện chính là yếu tố thúc đẩy việc khai thác, sử dụng tri thức số một cách hiệu quả. Vì vậy, nhằm làm rõ các vấn đề liên quan đến khả năng tiếp cận thông tin bằng ngôn ngữ ký hiệu, trong chương này tôi tập trung nghiên cứu những nội dung lý luận chung về ngôn ngữ ký hiệu và các phương pháp nhận dạng, với các nội dung chính như sau:

1.1. Các khái niệm liên quan đến nhận dạng ngôn ngữ ký hiệu

1.1.1. Khái niệm ngôn ngữ ký hiệu

Ngôn ngữ ký hiệu là một hệ thống ngôn ngữ tự nhiên được hình thành và phát triển trong cộng đồng người khiếm thính, sử dụng các chuyển động của bàn tay, biểu cảm khuôn mặt và tư thế cơ thể để truyền đạt thông tin. Khác với quan niệm phổ biến cho rằng ngôn ngữ ký hiệu chỉ là hình thức minh họa cho ngôn ngữ nói, thực tế đây là một ngôn ngữ độc lập, có cấu trúc ngữ pháp và từ vựng riêng [1, tr.166].

Các nghiên cứu ngôn ngữ học đã chỉ ra rằng ngôn ngữ ký hiệu mang đầy đủ đặc trưng của một ngôn ngữ tự nhiên như tính quy ước, tính hệ thống, khả năng sáng tạo và năng lực biểu đạt các khái niệm trừu tượng. Mỗi ký hiệu thường được cấu thành từ nhiều thành tố kết hợp đồng thời, bao gồm hình dạng bàn tay, vị trí thực hiện, hướng lòng bàn tay, chuyển động và các biểu cảm đi kèm.

Một đặc điểm quan trọng của ngôn ngữ ký hiệu là **không mang tính phổ quát**. Trên thế giới hiện tồn tại hàng trăm ngôn ngữ ký hiệu khác nhau, mỗi ngôn ngữ gắn với một cộng đồng và bối cảnh văn hóa cụ thể. Do đó, việc nghiên cứu và phát triển các hệ thống công nghệ liên quan đến ngôn ngữ ký hiệu cần được thực hiện riêng cho từng ngôn ngữ, trong đó có Ngôn ngữ ký hiệu Việt Nam.

1.1.2. Khái niệm nhận dạng ngôn ngữ ký hiệu

Nhận dạng ngôn ngữ ký hiệu là quá trình sử dụng các kỹ thuật thị giác máy tính và học máy nhằm tự động phân tích, xử lý và chuyển đổi các cử chỉ tay, chuyển động và biểu cảm trong ngôn ngữ ký hiệu thành dạng thông tin có thể hiểu và xử lý được bởi máy tính, chẳng hạn như văn bản, lệnh điều khiển hoặc tín hiệu truy vấn[4, tr.122].

Quá trình nhận dạng thường bao gồm các bước chính như thu nhận dữ liệu hình ảnh hoặc video, phát hiện và theo dõi bàn tay, trích xuất đặc trưng và áp dụng các mô hình học máy để phân loại hoặc dự đoán ký hiệu tương ứng [5, tr.81]. Trong những năm gần đây, sự phát triển của các mô hình học sâu đã giúp nâng cao đáng kể độ chính xác và khả năng tổng quát hóa của các hệ thống nhận dạng ngôn ngữ ký hiệu.

Tuy nhiên, việc nhận dạng ngôn ngữ ký hiệu trong môi trường thực tế vẫn gặp nhiều thách thức do sự đa dạng của người dùng, điều kiện ánh sáng, nền cảnh và sự khác biệt trong cách thực hiện ký hiệu. Điều này đặt ra yêu cầu cần xây dựng các mô hình và bộ dữ liệu phù hợp với bối cảnh sử dụng cụ thể.

1.1.3. Nhận dạng ngôn ngữ ký hiệu trong hệ thống thư viện số

Trong bối cảnh thư viện số, nhận dạng ngôn ngữ ký hiệu không chỉ dừng lại ở việc chuyển đổi ký hiệu thành văn bản, mà còn hướng tới việc hỗ trợ người khiếm thính tiếp cận và tra cứu thông tin một cách thuận tiện. Hệ thống nhận dạng cần đảm bảo khả năng hoạt động ổn định theo thời gian thực, độ chính xác đủ cao và dễ dàng tích hợp vào các nền tảng đang vận hành.

Việc ứng dụng nhận dạng ngôn ngữ ký hiệu vào thư viện số giúp giảm bớt rào cản trong quá trình tiếp cận tri thức, đồng thời góp phần nâng cao tính bao trùm và khả năng tiếp cận của các hệ thống thông tin hiện đại. Đây cũng chính là cơ sở để đề án tập trung nghiên cứu và phát triển hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam ứng dụng trong thư viện số.

1.1.4. Khái niệm về học máy (Machine Learning)

Học máy (Machine Learning) là một nhánh quan trọng của trí tuệ nhân tạo (Artificial Intelligence – AI), tập trung vào việc xây dựng các mô hình và thuật toán cho phép máy tính tự động học hỏi từ dữ liệu, nhận diện các mẫu (patterns) và đưa ra dự đoán hoặc quyết định mà không cần lập trình một cách tường minh cho từng trường hợp cụ thể [14, tr.12].

Trong học máy, mô hình được huấn luyện trên tập dữ liệu đầu vào nhằm tối ưu một hàm mục tiêu, từ đó có khả năng tổng quát hóa và áp dụng cho các dữ liệu chưa từng xuất hiện trước đó. Các phương pháp học máy thường được chia thành ba nhóm chính: học có giám sát, học không giám sát và học tăng cường. Trong đó, học có giám sát được sử dụng phổ biến trong các bài toán phân loại và nhận dạng, bao gồm bài toán nhận dạng ngôn ngữ ký hiệu.

Trong đề án này, học máy đóng vai trò cốt lõi trong việc xây dựng mô hình nhận dạng các ký hiệu tay, thông qua việc học mối quan hệ giữa dữ liệu đầu vào (chuỗi điểm mốc bàn tay theo thời gian) và nhãn đầu ra (ký hiệu hoặc cụm từ tương ứng).

1.1.5. Khái niệm về học sâu (Deep Learning)

Học sâu (Deep Learning) là một nhánh con của học máy, sử dụng các mạng nơ-ron nhân tạo nhiều lớp (Deep Neural Networks) [3, tr.236] để học và trích xuất đặc trưng từ dữ liệu ở nhiều mức độ trừu tượng khác nhau. Khác với các phương pháp học máy truyền thống, học sâu có khả năng tự động học đặc trưng từ dữ liệu thô mà không cần thiết kế thủ công.

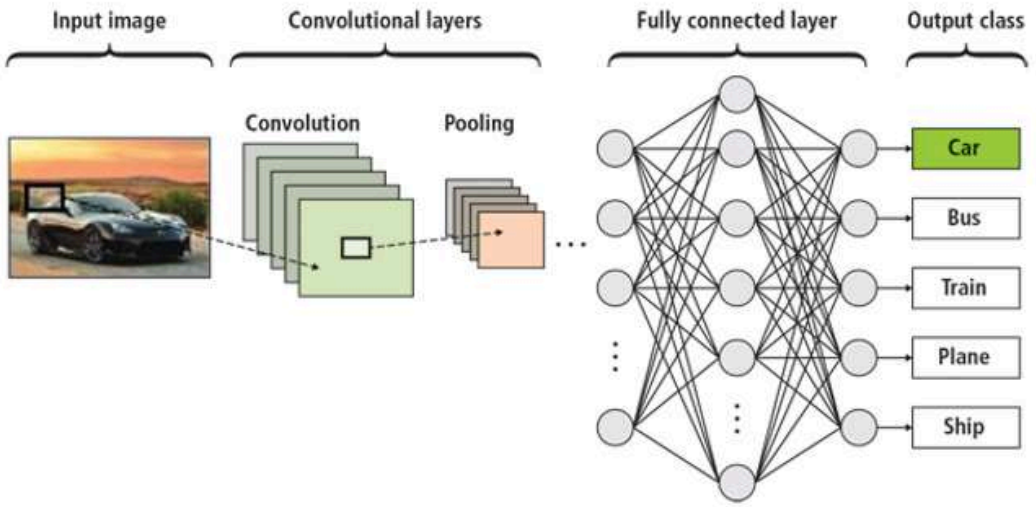
Các mô hình học sâu như mạng tích chập (CNN) [4, tr.12], mạng hồi tiếp (RNN), LSTM hay Transformer [4, tr.16] đã chứng minh hiệu quả vượt trội trong các bài toán xử lý ảnh, video và chuỗi thời gian. Đặc biệt, trong bài toán nhận dạng ngôn ngữ ký hiệu – vốn có tính chất vừa không gian (hình dạng bàn tay) vừa thời gian (chuyển động), học sâu cho phép mô hình hóa hiệu quả các đặc trưng phức tạp này.

Trong đề án, mô hình CNN–LSTM được lựa chọn nhằm kết hợp ưu điểm của mạng tích chập trong trích xuất đặc trưng không gian và LSTM trong học các mẫu biến đổi theo thời gian của chuỗi ký hiệu.

1.1.5. Khái niệm về thị giác máy tính (Computer Vision)

Thị giác máy tính (Computer Vision) là lĩnh vực nghiên cứu giúp máy tính có khả năng "nhìn" và hiểu được nội dung của hình ảnh hoặc video, tương tự như cách con người cảm nhận thế giới bằng thị giác. Các bài toán phổ biến của thị giác máy tính bao gồm nhận dạng đối tượng, phát hiện khuôn mặt, theo dõi chuyển động, phân tích hành vi và nhận dạng cử chỉ.

Trong bối cảnh nhận dạng ngôn ngữ ký hiệu, thị giác máy tính đóng vai trò xử lý và phân tích hình ảnh bàn tay người dùng, trích xuất thông tin hình học và chuyển động làm đầu vào cho mô hình học máy. Việc áp dụng thị giác máy tính giúp giảm sự phụ thuộc vào các thiết bị chuyên dụng, cho phép hệ thống hoạt động với camera thông thường, phù hợp với môi trường thư viện số.



Hình 1.1. Tổng quan mô hình thị giác máy tính

1.1.6. Khái niệm về nhận dạng cử chỉ và ngôn ngữ ký hiệu

Nhận dạng cử chỉ (Gesture Recognition) là bài toán xác định và phân loại các chuyển động hoặc tư thế của bàn tay, cánh tay hoặc cơ thể người từ dữ liệu hình ảnh hoặc video. Nhận dạng ngôn ngữ ký hiệu là một dạng đặc thù của nhận dạng cử chỉ, trong đó các cử chỉ mang ý nghĩa ngôn ngữ và tuân theo các quy tắc nhất định.
Ngôn ngữ ký hiệu Việt Nam có đặc điểm riêng về hình dạng bàn tay, hướng, vị trí và chuyển động, do đó việc nhận dạng đòi hỏi mô hình phải học được cả đặc trưng không gian lẫn đặc trưng thời gian. Đây là một bài toán phức tạp, chịu ảnh hưởng bởi nhiều yếu tố như điều kiện ánh sáng, nền cảnh, góc quay camera và sự khác biệt giữa các cá nhân thực hiện ký hiệu [2, tr.466].

1.2. Tổng quan nghiên cứu trong nước và ngoài nước

1.2.1. Các nghiên cứu trên thế giới

Trên thế giới, nhận dạng ngôn ngữ ký hiệu là một hướng nghiên cứu quan trọng trong lĩnh vực thị giác máy tính và trí tuệ nhân tạo. Các nghiên cứu ban đầu chủ yếu dựa trên phương pháp trích xuất đặc trưng thủ công kết hợp với các thuật toán phân loại truyền thống. Những năm gần đây, các mô hình học sâu như CNN [2, tr.166], RNN [3, tr.11] và Transformer [11, tr.2] đã được áp dụng rộng rãi, giúp nâng cao độ chính xác và khả năng xử lý ký hiệu động.

Một số hệ thống đã được triển khai thử nghiệm trong các ứng dụng hỗ trợ giáo dục, giao tiếp và tra cứu thông tin, cho thấy tiềm năng lớn của công nghệ này trong việc hỗ trợ người khiếm thính.



Hình 1.2. Thủ ngữ - Ngôn ngữ ký hiệu tay theo chữ cái latin

(Nguồn: www.nki.vn)

1.2.2. Các nghiên cứu tại Việt Nam

Tại Việt Nam, các nghiên cứu về nhận dạng ngôn ngữ ký hiệu mới được quan tâm trong những năm gần đây. Phần lớn các công trình tập trung vào xây dựng bộ dữ liệu ngôn ngữ ký hiệu Việt Nam và thử nghiệm nhận dạng các ký hiệu đơn lẻ trong môi trường kiểm soát. Một số nghiên cứu đạt được độ chính xác tương đối cao, tuy nhiên quy mô dữ liệu còn hạn chế và chưa phản ánh đầy đủ điều kiện sử dụng thực tế.

Đến nay, chưa có nhiều nghiên cứu triển khai hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam tích hợp vào các hệ thống thông tin đang vận hành như thư viện số, cổng dịch vụ công hay nền tảng giáo dục trực tuyến.

1.2.3. Nhận xét và khoảng trống nghiên cứu

Từ tổng quan các nghiên cứu trong và ngoài nước có thể nhận thấy một số hạn chế còn tồn tại. Thứ nhất, nhiều nghiên cứu mới chỉ tập trung vào ký hiệu đơn lẻ mà chưa mở rộng sang các cụm từ phục vụ tác vụ tra cứu thông tin. Thứ hai, các mô hình thường được đánh giá trong môi trường phòng thí nghiệm, chưa kiểm

chứng đầy đủ trong điều kiện thực tế. Thứ ba, việc tích hợp công nghệ nhận dạng ngôn ngữ ký hiệu vào các hệ thống thư viện số vẫn còn rất ít được đề cập.

Những khoảng trống này chính là cơ sở để đề án tiếp tục nghiên cứu và phát triển giải pháp phù hợp với bối cảnh ứng dụng thực tế.

1.3. Định hướng nghiên cứu của đề án

Trên cơ sở hệ thống hóa các vấn đề lý luận và tổng quan các nghiên cứu trong và ngoài nước, đề án được định hướng theo hướng nghiên cứu ứng dụng, kết hợp chặt chẽ giữa lý thuyết và triển khai thực tế. Trọng tâm của đề án không chỉ dừng lại ở việc xây dựng mô hình nhận dạng ngôn ngữ ký hiệu Việt Nam mà còn hướng tới việc tích hợp mô hình này vào một hệ thống thư viện số đang vận hành, nhằm giải quyết nhu cầu tiếp cận thông tin của người khiếm thính trong bối cảnh thực tiễn.

Cụ thể, đề án tập trung xây dựng bộ dữ liệu ngôn ngữ ký hiệu Việt Nam phục vụ trực tiếp cho tác vụ tìm kiếm thư viện số, bao gồm các ký hiệu và cụm từ thường xuyên được sử dụng trong hoạt động tra cứu tài liệu. Dữ liệu được thu thập từ người dùng thực tế trong nhiều điều kiện khác nhau về ánh sáng, nền cảnh và góc quay, nhằm phản ánh sát môi trường sử dụng và nâng cao khả năng tổng quát của mô hình.

Bên cạnh đó, đề án định hướng nghiên cứu và lựa chọn các phương pháp học máy phù hợp với đặc thù dữ liệu chuỗi thời gian của ngôn ngữ ký hiệu, ưu tiên các mô hình có khả năng cân bằng giữa độ chính xác và hiệu năng xử lý để đáp ứng yêu cầu nhận dạng theo thời gian thực. Việc đánh giá mô hình không chỉ dựa trên các chỉ số kỹ thuật mà còn xem xét khả năng triển khai, tích hợp và vận hành ổn định trong hệ thống thực tế.

Ngoài khía cạnh kỹ thuật, đề án cũng chú trọng yếu tố người dùng, thông qua việc thiết kế giao diện tương tác thân thiện, dễ sử dụng và phù hợp với đặc điểm của cộng đồng người khiếm thính. Việc tích hợp hệ thống nhận dạng ngôn ngữ ký hiệu vào nền tảng thư viện số thuvienso.digilib.vn được xem là bước kiểm chứng quan

trọng, nhằm đánh giá tính khả thi, hiệu quả và tiềm năng mở rộng của giải pháp trong thực tế.

Thông qua định hướng nghiên cứu này, đề án kỳ vọng đóng góp cả về mặt học thuật lẫn ứng dụng, góp phần hoàn thiện các giải pháp công nghệ hỗ trợ người khiếm thính và thúc đẩy khả năng tiếp cận thông tin trong các hệ thống thư viện số tại Việt Nam.

CHƯƠNG 2:
PHÂN TÍCH, THIẾT KẾ VÀ PHÁT TRIỂN HỆ THỐNG NHẬN DẠNG NGÔN NGỮ KÝ HIỆU ỨNG DỤNG TRONG THƯ VIỆN SỐ

2.1. Phân tích yêu cầu hệ thống

2.1.1. Phân tích đối tượng sử dụng hệ thống

Hệ thống thư viện số thuvienso.digilib.vn được thiết kế phục vụ đa dạng nhóm người dùng, trong đó đối tượng sử dụng chính là học sinh, sinh viên tại các cơ sở giáo dục, đồng thời mở rộng cho cán bộ, giáo viên, giảng viên và người học trong các đơn vị, sở, ngành và doanh nghiệp tham gia hệ sinh thái thư viện số.

Theo mô hình triển khai hiện tại, hệ thống được tổ chức theo cấu trúc site – đơn vị – cơ quan quản lý**,** cho phép vận hành linh hoạt ở nhiều cấp khác nhau như: thư viện trường học, thư viện sở/ban/ngành, thư viện doanh nghiệp và các thư viện dùng chung trên phạm vi địa phương hoặc toàn ngành. Mỗi site có tập người dùng riêng, dữ liệu riêng và cơ chế phân quyền phù hợp với đặc thù quản lý.

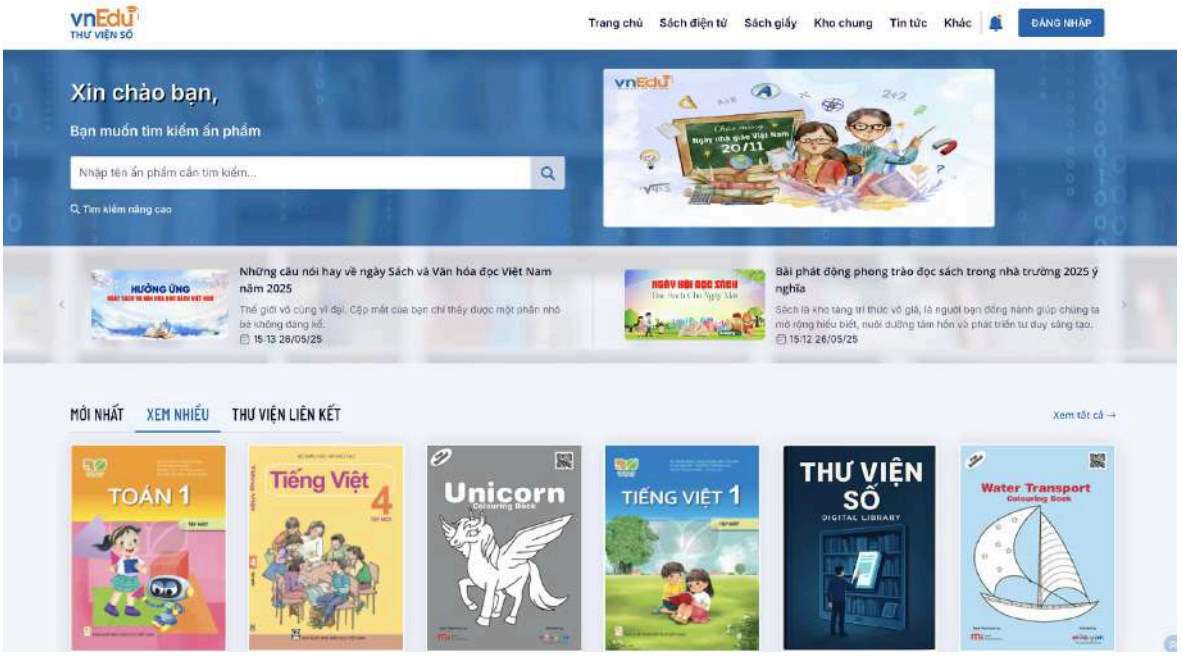
Trong bối cảnh đó, đề án này không thay đổi đối tượng sử dụng cốt lõi của hệ thống, mà tập trung mở rộng thêm một module chức năng hỗ trợ người khiếm thính – đặc biệt là học sinh, sinh viên khiếm thính đang học tập tại các cơ sở giáo dục hoặc tham gia các chương trình đào tạo nghề, đào tạo doanh nghiệp có sử dụng thư viện số.

Đối với nhóm người dùng khiếm thính, việc tra cứu tài liệu bằng văn bản tiếng Việt gặp nhiều rào cản do khác biệt về ngôn ngữ tiếp cận, hạn chế về chính tả và khó khăn trong việc sử dụng các từ khóa học thuật. Trong khi đó, hệ thống thư viện số hiện nay chủ yếu được thiết kế cho người nghe – nói, chưa tính đến nhu cầu tiếp cận bằng ngôn ngữ ký hiệu.

Vì vậy, module nhận dạng ngôn ngữ ký hiệu được đề xuất trong đề án đóng vai trò là lớp hỗ trợ bổ sung, cho phép người khiếm thính sử dụng ngôn ngữ ký hiệu Việt

Nam để thực hiện thao tác tìm kiếm tài liệu. Module này được tích hợp trực tiếp vào hệ thống hiện có, hoạt động song song với chức năng tìm kiếm truyền thống, đảm bảo không ảnh hưởng đến các nhóm người dùng khác.

Ngoài người học, module cũng có thể hỗ trợ giáo viên, cán bộ đào tạo, nhân viên thư viện trong quá trình hướng dẫn người khiếm thính tiếp cận tài nguyên số, đặc biệt trong các môi trường giáo dục hòa nhập và đào tạo nghề. Điều này góp phần mở rộng phạm vi phục vụ của thư viện số, hướng tới mục tiêu tiếp cận bình đẳng và bao trùm (inclusive access) trong chuyển đổi số giáo dục.



Hình 2.1. Hình ảnh trang chủ hệ thống thư viện số do VNPT-IT phát triển

(Nguồn: vnedustore.digilib.vn)

Ngoài nhóm đối tượng trong lĩnh vực giáo dục, đề án cũng định hướng mở rộng hệ thống cho khối doanh nghiệp. Trong bối cảnh doanh nghiệp ngày càng chú trọng đào tạo nội bộ và xây dựng kho tri thức số, việc tích hợp thư viện số cùng các công cụ hỗ trợ tiếp cận thông tin sẽ góp phần nâng cao hiệu quả học tập, đào tạo và chia sẻ tri thức trong tổ chức.

Đặc biệt, một điểm nhấn quan trọng của hệ thống là module hỗ trợ người khiếm thính, trong đó đối tượng phục vụ bao gồm học sinh, sinh viên và người lao động câm điếc có nhu cầu tiếp cận tài nguyên thư viện. Module này cho phép người

dùng sử dụng ngôn ngữ ký hiệu để tương tác với hệ thống, góp phần giảm rào cản giao tiếp, nâng cao khả năng tiếp cận tri thức và thúc đẩy giáo dục hòa nhập.

2.1.2. Phân tích bối cảnh và hệ thống thư viện số hiện tại

Trong những năm gần đây, cùng với quá trình chuyển đổi số trong lĩnh vực giáo dục và quản lý tri thức, các hệ thống thư viện số ngày càng đóng vai trò quan trọng trong việc lưu trữ, quản lý và cung cấp tài nguyên học tập, nghiên cứu cho người dùng. Thư viện số không chỉ là nơi tra cứu tài liệu mà còn trở thành nền tảng hỗ trợ học tập trực tuyến, đào tạo từ xa và chia sẻ tri thức trong nhiều mô hình tổ chức khác nhau như nhà trường, cơ quan quản lý và doanh nghiệp.

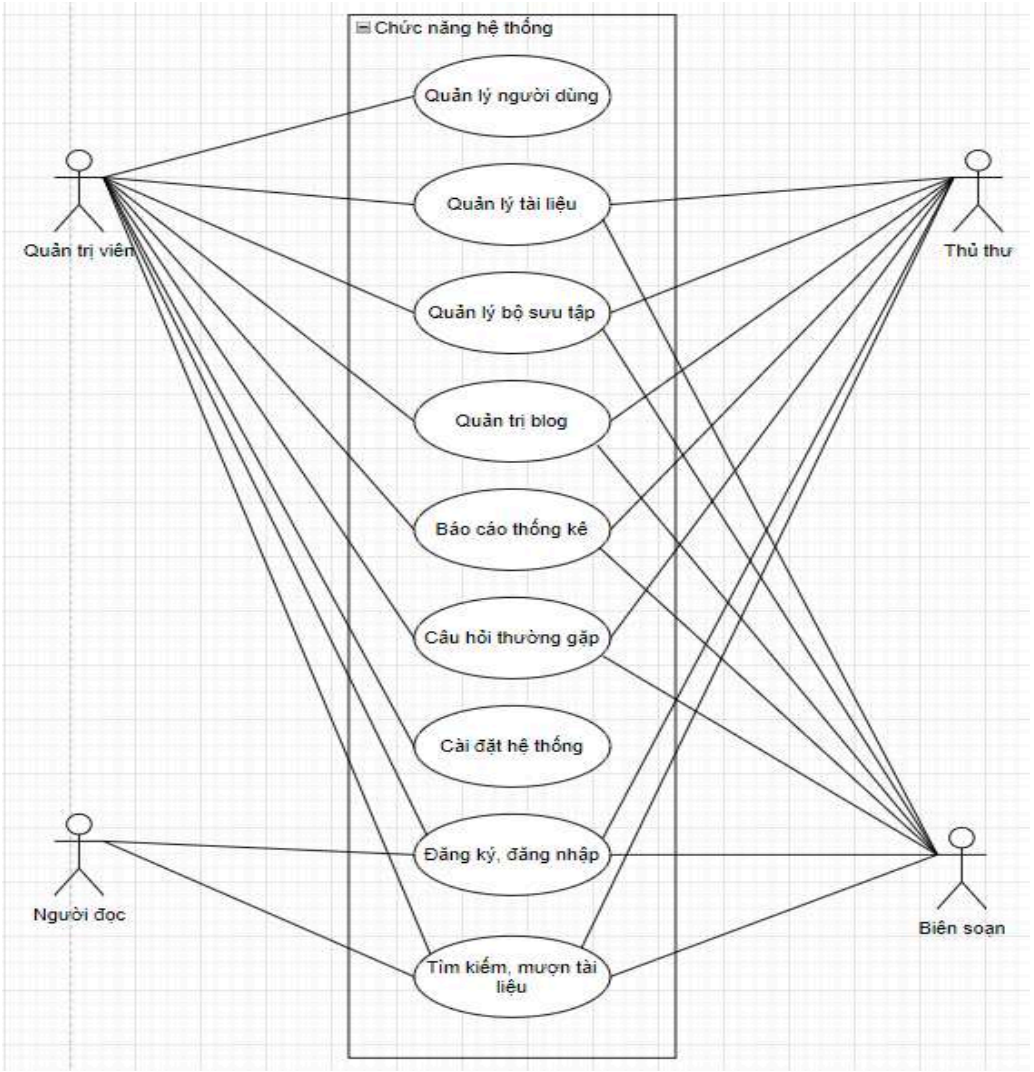
Hệ thống thư viện số hiện tại được triển khai theo mô hình tập trung, hỗ trợ nhiều site và đơn vị sử dụng chung một nền tảng. Mỗi site có thể đại diện cho một trường học, trung tâm đào tạo, đơn vị trực thuộc hoặc một tổ chức độc lập, trong khi cấp quản lý cao hơn (như sở, ngành) có thể thực hiện vai trò giám sát, tổng hợp và quản trị dữ liệu. Mô hình này giúp tối ưu việc vận hành, giảm chi phí triển khai và đảm bảo khả năng mở rộng khi số lượng người dùng và tài nguyên ngày càng tăng.
Về mặt chức năng, hệ thống thư viện số hiện tại đã đáp ứng tốt các nhu cầu cơ bản như: quản lý tài nguyên số, tìm kiếm và tra cứu tài liệu, phân quyền người dùng, thống kê sử dụng và hỗ trợ truy cập trên nhiều thiết bị khác nhau. Giao diện được thiết kế thân thiện, phù hợp với học sinh, sinh viên và cán bộ quản lý, đồng thời cho phép tích hợp các dịch vụ mở rộng thông qua cơ chế API.

Tuy nhiên, qua quá trình phân tích và khảo sát thực tế, có thể nhận thấy rằng hệ thống hiện tại **chưa có các chức năng hỗ trợ chuyên biệt cho nhóm người dùng khiếm thính**, đặc biệt là trong khâu tương tác và tìm kiếm thông tin. Người dùng câm điếc vẫn phải sử dụng các hình thức nhập liệu truyền thống như bàn phím hoặc chuột, trong khi ngôn ngữ ký hiệu – phương thức giao tiếp tự nhiên của họ – chưa được khai thác trong hệ thống.

Bên cạnh đó, hệ thống thư viện số hiện tại chủ yếu phục vụ người dùng phổ thông, chưa tích hợp các công nghệ thông minh nhằm hỗ trợ nhận dạng hành vi, cử chỉ hoặc tương tác trực quan. Điều này đặt ra yêu cầu cần mở rộng hệ thống theo

hướng tích hợp các dịch vụ xử lý thông minh, trong đó có công nghệ nhận dạng ngôn ngữ ký hiệu, nhưng vẫn phải đảm bảo không làm ảnh hưởng đến kiến trúc tổng thể và hoạt động ổn định của hệ thống.

Từ bối cảnh trên, việc nghiên cứu và phát triển module nhận dạng ngôn ngữ ký hiệu tích hợp vào hệ thống thư viện số là cần thiết và phù hợp với định hướng phát triển hiện tại. Giải pháp này không chỉ nâng cao khả năng tiếp cận thông tin cho người khiếm thính mà còn góp phần hoàn thiện hệ sinh thái thư viện số theo hướng hiện đại, nhân văn và bao trùm hơn.



Hình 2.2. Hình ảnh mô tả chức năng của người dùng với hệ thống thư viện số

2.1.3. Xác định yêu cầu chức năng của hệ thống

Dựa trên kết quả phân tích bối cảnh sử dụng, đặc điểm người dùng và hệ thống thư viện số hiện tại, việc xác định đầy đủ các yêu cầu chức năng là cơ sở quan trọng để thiết kế và triển khai giải pháp nhận dạng ngôn ngữ ký hiệu một cách hiệu quả. Các yêu cầu chức năng của hệ thống được xác định theo hai nhóm chính: chức năng của hệ thống thư viện số hiện hữu và chức năng của module nhận dạng ngôn ngữ ký hiệu tích hợp.

a) Yêu cầu chức năng của hệ thống thư viện số

Hệ thống thư viện số cần tiếp tục đảm bảo đầy đủ các chức năng cốt lõi nhằm phục vụ hoạt động học tập, giảng dạy và nghiên cứu của học sinh, sinh viên, cán bộ và các đơn vị sử dụng. Các chức năng chính bao gồm:

- **Quản lý người dùng và phân quyền**: Hệ thống cho phép quản lý nhiều nhóm người dùng khác nhau như quản trị viên hệ thống, cán bộ thư viện, giáo viên, học sinh, sinh viên và người dùng doanh nghiệp. Mỗi nhóm được phân quyền truy cập và sử dụng tài nguyên phù hợp với vai trò và cấp quản lý (site, đơn vị, sở/ngành).
- **Quản lý tài nguyên số**: Cho phép thêm mới, chỉnh sửa, phân loại và lưu trữ các tài liệu số như sách điện tử, giáo trình, bài giảng, luận văn, video và các tài liệu đa phương tiện khác. Hệ thống hỗ trợ gán metadata, từ khóa và phân loại theo chủ đề nhằm phục vụ cho việc tìm kiếm.
- **Tìm kiếm và tra cứu thông tin**: Cung cấp chức năng tìm kiếm tài liệu theo nhiều tiêu chí như tên tài liệu, tác giả, chủ đề, năm xuất bản và từ khóa. Hệ thống hỗ trợ tìm kiếm toàn văn và lọc kết quả để người dùng dễ dàng tiếp cận tài liệu cần thiết.
- **Theo dõi và thống kê sử dụng**: Ghi nhận lịch sử truy cập, lượt xem và tải tài liệu của người dùng; đồng thời cung cấp các báo cáo thống kê phục vụ

công tác quản lý, đánh giá hiệu quả khai thác tài nguyên tại từng site và đơn vị.

b) Yêu cầu chức năng của module nhận dạng ngôn ngữ ký hiệu

Nhằm tăng cường khả năng tiếp cận thông tin cho người khiếm thính, đề án đề xuất phát triển một module nhận dạng ngôn ngữ ký hiệu được tích hợp trực tiếp vào hệ thống thư viện số. Module này cần đáp ứng các yêu cầu chức năng sau:

- **Nhận dạng ngôn ngữ ký hiệu theo thời gian thực**: Hệ thống cho phép người dùng khiếm thính sử dụng camera để thực hiện các ký hiệu tay, từ đó nhận dạng và chuyển đổi ký hiệu thành văn bản hoặc lệnh tìm kiếm tương ứng trong thư viện số.
- **Hỗ trợ tìm kiếm tài liệu bằng ký hiệu**: Kết quả nhận dạng từ ngôn ngữ ký hiệu được sử dụng như đầu vào cho chức năng tìm kiếm, giúp người dùng khiếm thính có thể tra cứu tài liệu mà không cần nhập liệu bằng bàn phím.
- **Xử lý và phản hồi kết quả**: Sau khi nhận dạng, hệ thống hiển thị kết quả tìm kiếm dưới dạng danh sách tài liệu kèm thông tin mô tả, đồng thời hỗ trợ xem trước nội dung để người dùng lựa chọn tài liệu phù hợp.
- **Hoạt động độc lập và tích hợp linh hoạt**: Module nhận dạng được triển khai như một dịch vụ độc lập, giao tiếp với hệ thống thư viện số thông qua API. Cách tiếp cận này giúp đảm bảo không ảnh hưởng đến hoạt động hiện tại của hệ thống và thuận tiện cho việc mở rộng trong tương lai.
- **Quản lý và ghi nhận dữ liệu sử dụng**: Hệ thống ghi nhận các truy vấn tìm kiếm bằng ký hiệu nhằm phục vụ việc phân tích, cải tiến chất lượng nhận dạng và nâng cao trải nghiệm người dùng.

c) Yêu cầu chức năng mở rộng

Ngoài các chức năng cơ bản, hệ thống cần đảm bảo khả năng mở rộng để đáp ứng nhu cầu sử dụng trong tương lai, bao gồm:

- Mở rộng tập ký hiệu và cụm từ phù hợp với ngữ cảnh thư viện.
- Hỗ trợ nhiều nhóm người dùng và nhiều mô hình triển khai khác nhau như trường học, cơ quan quản lý và doanh nghiệp.

- Cho phép tích hợp thêm các module hỗ trợ tiếp cận khác cho người khuyết tật.

Việc xác định đầy đủ các yêu cầu chức năng trên là cơ sở quan trọng cho các bước thiết kế kiến trúc hệ thống, xây dựng mô hình nhận dạng và triển khai giải pháp trong các mục tiếp theo của đề án.

2.1.4. Xác định yêu cầu phi chức năng

Bên cạnh các yêu cầu chức năng, hệ thống thư viện số tích hợp module nhận dạng ngôn ngữ ký hiệu cần đáp ứng đầy đủ các yêu cầu phi chức năng nhằm đảm bảo khả năng vận hành ổn định, hiệu quả và phù hợp với môi trường triển khai thực tế. Các yêu cầu phi chức năng được xác định bao gồm các khía cạnh về hiệu năng, độ chính xác, khả năng mở rộng, bảo mật, khả năng sử dụng và tính tương thích.

a) Yêu cầu về hiệu năng

Hệ thống cần đảm bảo khả năng xử lý nhanh và ổn định trong điều kiện có nhiều người dùng truy cập đồng thời, đặc biệt tại các site trường học và đơn vị quản lý có số lượng lớn học sinh, sinh viên. Đối với module nhận dạng ngôn ngữ ký hiệu, thời gian phản hồi từ lúc người dùng thực hiện ký hiệu đến khi hiển thị kết quả tìm kiếm cần đủ nhanh để không gây gián đoạn trải nghiệm, phù hợp với tương tác thời gian thực. Hệ thống phải duy trì hiệu năng ổn định ngay cả khi khối lượng dữ liệu và số lượng truy vấn tăng lên.

b) Yêu cầu về độ chính xác và độ tin cậy

Mô hình nhận dạng ngôn ngữ ký hiệu cần đạt độ chính xác ở mức cao và ổn định trong nhiều điều kiện sử dụng khác nhau như thay đổi ánh sáng, nền cảnh và góc quay camera. Kết quả nhận dạng phải đủ tin cậy để phục vụ tác vụ tìm kiếm tài liệu, hạn chế tối đa việc nhận dạng sai gây ảnh hưởng đến trải nghiệm của người dùng khiếm thính. Hệ thống cần có cơ chế xử lý lỗi và thông báo phù hợp trong trường hợp nhận dạng không thành công hoặc dữ liệu đầu vào không đảm bảo chất lượng.

c) Yêu cầu về khả năng mở rộng và bảo trì

Hệ thống phải được thiết kế theo hướng module hóa, cho phép mở rộng thêm các chức năng mới mà không ảnh hưởng đến các thành phần đang hoạt động. Module nhận dạng ngôn ngữ ký hiệu cần có khả năng mở rộng tập ký hiệu, cụm từ và mô hình học máy trong tương lai. Bên cạnh đó, việc bảo trì, nâng cấp và thay thế mô hình nhận dạng cần được thực hiện thuận tiện, không làm gián đoạn hoạt động chung của thư viện số.

d) Yêu cầu về bảo mật và an toàn dữ liệu

Hệ thống cần tuân thủ các yêu cầu về bảo mật thông tin, đặc biệt là dữ liệu người dùng và dữ liệu hình ảnh, video thu thập trong quá trình nhận dạng ngôn ngữ ký hiệu. Việc thu thập và xử lý dữ liệu phải được sự đồng ý của người dùng và tuân thủ các quy định hiện hành về bảo vệ dữ liệu cá nhân. Các cơ chế xác thực, phân quyền và kiểm soát truy cập cần được áp dụng thống nhất trên toàn hệ thống, đảm bảo an toàn cho dữ liệu tại từng site và đơn vị.

e) Yêu cầu về khả năng sử dụng và tiếp cận

Giao diện hệ thống cần thân thiện, dễ sử dụng và phù hợp với nhiều nhóm người dùng khác nhau, đặc biệt là học sinh, sinh viên và người khiếm thính. Module nhận dạng ngôn ngữ ký hiệu phải được thiết kế đơn giản, dễ thao tác, giảm tối đa các bước trung gian trong quá trình tìm kiếm tài liệu. Hệ thống cần đảm bảo các nguyên tắc tiếp cận số, hỗ trợ người dùng khuyết tật tiếp cận thông tin một cách thuận tiện và bình đẳng.

f) Yêu cầu về tính tương thích và triển khai

Hệ thống phải tương thích với hạ tầng thư viện số hiện có, hoạt động ổn định trên các trình duyệt phổ biến và nhiều loại thiết bị như máy tính để bàn, laptop và thiết bị di động. Module nhận dạng ngôn ngữ ký hiệu cần được triển khai linh hoạt dưới dạng dịch vụ độc lập, có thể tích hợp thông qua API mà không yêu cầu thay đổi lớn về kiến trúc hệ thống hiện tại.

2.2. Phân tích bài toán nhận dạng ngôn ngữ ký hiệu

2.2.1. Mô tả bài toán

Bài toán nhận dạng ngôn ngữ ký hiệu trong đề án này được đặt ra nhằm chuyển đổi các cử chỉ tay của người khiếm thính, được ghi nhận thông qua thiết bị camera, thành thông tin dạng văn bản hoặc lệnh tìm kiếm có ý nghĩa, từ đó hỗ trợ người dùng tra cứu và tiếp cận tài nguyên trong thư viện số một cách thuận tiện hơn. Cụ thể, hệ thống cần nhận diện các ký hiệu trong ngôn ngữ ký hiệu Việt Nam (Vietnamese Sign Language – VSL) và ánh xạ chúng tới các từ khóa hoặc cụm từ tương ứng được sử dụng trong hoạt động tìm kiếm tài liệu.

Về mặt kỹ thuật, bài toán được mô hình hóa như một bài toán phân loại chuỗi thời gian trong lĩnh vực thị giác máy tính và học máy. Đầu vào của hệ thống là một chuỗi các khung hình video ghi lại quá trình thực hiện ký hiệu tay của người dùng. Thông qua quá trình xử lý hình ảnh và trích xuất đặc trưng, mỗi ký hiệu được biểu diễn dưới dạng chuỗi điểm đặc trưng không gian–thời gian, phản ánh chuyển động và cấu hình của bàn tay theo thời gian. Đầu ra của hệ thống là nhãn ký hiệu tương ứng, bao gồm các ký hiệu đơn lẻ và một số cụm từ thường dùng trong tra cứu thư viện số.

Khác với các bài toán nhận dạng ký hiệu trong môi trường phòng thí nghiệm, bài toán trong đề án này hướng tới bối cảnh sử dụng thực tế, nơi người dùng là học sinh, sinh viên và các nhóm người sử dụng thư viện số tại các đơn vị, sở, ngành và doanh nghiệp. Do đó, hệ thống cần có khả năng hoạt động ổn định trong điều kiện đa dạng về ánh sáng, nền cảnh, góc quay camera và sự khác biệt giữa các người ký hiệu. Đồng thời, thời gian phản hồi của hệ thống phải đủ nhanh để đảm bảo trải nghiệm tương tác liên tục, phù hợp với yêu cầu vận hành của một nền tảng thư viện số đang hoạt động.

Phạm vi bài toán trong đề án được giới hạn ở việc nhận dạng một tập ký hiệu và cụm từ ngôn ngữ ký hiệu Việt Nam có liên quan trực tiếp đến hoạt động tìm kiếm và truy cập tài nguyên thư viện, thay vì xử lý toàn bộ ngôn ngữ ký hiệu ở mức hội thoại liên tục. Cách tiếp cận này giúp cân bằng giữa độ phức tạp kỹ thuật, độ

chính xác của mô hình và khả năng triển khai thực tế trên hệ thống thuvienso.digilib.vn, đồng thời tạo tiền đề để mở rộng và hoàn thiện hệ thống trong các giai đoạn phát triển tiếp theo.

Cơ sở lý thuyết giải quyết bài toán hỗ trợ người khuyết tật trong môi trường số:

Trong bối cảnh chuyển đổi số diễn ra mạnh mẽ, việc bảo đảm khả năng tiếp cận thông tin cho người khuyết tật trở thành một yêu cầu quan trọng, không chỉ mang ý nghĩa nhân văn mà còn là trách nhiệm xã hội và nghĩa vụ pháp lý của các tổ chức cung cấp dịch vụ công. Đối với người khiếm thính hoặc người câm điếc, rào cản lớn nhất trong việc tiếp cận dịch vụ số không nằm ở việc thiếu thiết bị hay kết nối Internet, mà chủ yếu đến từ hạn chế về giao tiếp ngôn ngữ. Trong khi phần lớn các hệ thống số hiện nay vẫn dựa vào văn bản và giọng nói, nhóm người dùng này lại có nhu cầu giao tiếp tự nhiên bằng ngôn ngữ ký hiệu – một hệ thống ngôn ngữ độc lập với cấu trúc và ngữ pháp riêng.

Về mặt lý thuyết, vấn đề hỗ trợ người khuyết tật trong môi trường số được đặt trong khung tiếp cận của thiết kế truy cập (accessibility design) và thiết kế toàn diện (universal design). Theo đó, hệ thống công nghệ cần được xây dựng theo hướng tạo điều kiện để mọi nhóm người dùng, bao gồm cả nhóm yếu thế, đều có thể sử dụng dịch vụ một cách bình đẳng và thuận tiện. Trong lĩnh vực web và dịch vụ số, bộ tiêu chuẩn được nhắc đến phổ biến nhất là WCAG[], đưa ra các nguyên tắc thiết kế nhằm bảo đảm nội dung có thể được nhận biết – vận hành – hiểu được – tương thích với nhiều đối tượng người dùng khác nhau. Tuy nhiên, đối với người khiếm thính, việc đáp ứng WCAG mới chỉ giải quyết được một phần, vì thách thức lớn hơn nằm ở kênh giao tiếp ngôn ngữ, đặc biệt trong các tác vụ cần tương tác nhanh như tìm kiếm thông tin.

Từ góc nhìn tương tác người – máy (Human–Computer Interaction), việc hỗ trợ người khiếm thính trong thư viện số có thể được xem là bài toán mở rộng kênh tương tác. Thay vì chỉ sử dụng bàn phím hoặc nhập văn bản, hệ thống cần bổ sung

một phương thức tương tác trực quan và tự nhiên hơn: tìm kiếm bằng ngôn ngữ ký hiệu. Đây là một dạng giao tiếp đa phương thức, trong đó người dùng sử dụng cử chỉ tay và biểu cảm cơ thể, còn hệ thống sử dụng camera và mô hình học máy để chuyển đổi tín hiệu đó thành văn bản hoặc từ khóa tìm kiếm. Mô hình tương tác này có ý nghĩa quan trọng vì nó giúp người dùng tiếp cận hệ thống theo cách gần với ngôn ngữ tự nhiên của họ nhất, từ đó giảm gánh nặng phải "dịch" sang văn bản tiếng Việt trong quá trình sử dụng.

Trên phương diện kỹ thuật, việc xây dựng hệ thống nhận dạng ngôn ngữ ký hiệu phục vụ người khuyết tật thường được đặt trong lĩnh vực thị giác máy tính (computer vision)[] và học máy/học sâu (machine learning/deep learning)[]. Bài toán cốt lõi là nhận dạng các ký hiệu từ dữ liệu hình ảnh hoặc video, sau đó ánh xạ kết quả nhận dạng sang nội dung có ý nghĩa (từ khóa hoặc câu truy vấn). Trong đề án này, dữ liệu đầu vào được biểu diễn dưới dạng chuỗi điểm mốc bàn tay (hand landmarks) theo thời gian. Đây là dạng dữ liệu chuỗi, vì mỗi ký hiệu thường được hình thành bởi chuyển động liên tục trong nhiều khung hình. Do đó, mô hình cần vừa học được đặc trưng không gian của bàn tay tại từng thời điểm, vừa học được mối quan hệ thời gian của toàn bộ chuỗi. Điều này dẫn tới việc lựa chọn các kiến trúc học sâu phù hợp như CNN[] để trích xuất đặc trưng không gian và LSTM[] để học mẫu thời gian, hoặc các kiến trúc hiện đại hơn như Transformer[].

Ngoài việc giải quyết bài toán nhận dạng, một khía cạnh quan trọng khác trong cơ sở lý thuyết là tính khả thi triển khai. Các hệ thống hỗ trợ người khuyết tật trong môi trường thực tế thường phải đáp ứng các yêu cầu nghiêm ngặt như: độ trễ thấp, khả năng hoạt động ổn định trên thiết bị phổ thông, dễ tích hợp vào nền tảng đang vận hành và đảm bảo quyền riêng tư dữ liệu. Do dữ liệu video và hình ảnh của người dùng có thể được xem là dữ liệu nhạy cảm, hệ thống cần có cơ chế xin phép, hạn chế lưu trữ, và chỉ gửi các đặc trưng cần thiết thay vì toàn bộ video thô, nhằm giảm rủi ro lộ lọt thông tin cá nhân.

Như vậy, xét trên cả khía cạnh xã hội, lý thuyết truy cập số và nền tảng kỹ thuật, bài toán nhận dạng ngôn ngữ ký hiệu ứng dụng trong thư viện số không chỉ là một bài toán thị giác máy tính đơn thuần. Đây là một hướng tiếp cận tổng hợp nhằm giải quyết rào cản giao tiếp cho người khiếm thính, góp phần nâng cao khả năng tiếp cận tri thức và dịch vụ số, đồng thời thể hiện tính ứng dụng thực tiễn của các mô hình học sâu trong việc hỗ trợ nhóm người dùng yếu thế.

2.2.1. Đặc thù của ngôn ngữ ký hiệu Việt Nam trong bài toán nhận dạng

Ngôn ngữ ký hiệu Việt Nam là một hệ thống ngôn ngữ tự nhiên được hình thành và phát triển trong cộng đồng người khiếm thính tại Việt Nam. Khác với ngôn ngữ nói, ngôn ngữ ký hiệu Việt Nam truyền tải thông tin thông qua sự kết hợp đồng thời của nhiều yếu tố như hình dạng bàn tay, vị trí thực hiện, hướng chuyển động, quỹ đạo di chuyển và các biểu cảm phi ngôn ngữ đi kèm. Những đặc điểm này tạo nên tính phong phú trong biểu đạt, nhưng đồng thời cũng làm tăng độ phức tạp cho bài toán nhận dạng tự động.

Một đặc trưng quan trọng của ngôn ngữ ký hiệu Việt Nam là **tính không tuyến tính theo thời gian**. Trong nhiều trường hợp, ý nghĩa của một ký hiệu không chỉ phụ thuộc vào từng khung hình riêng lẻ mà còn nằm ở mối quan hệ giữa các chuyển động liên tiếp của bàn tay trong suốt quá trình thực hiện ký hiệu. Điều này đòi hỏi mô hình nhận dạng phải có khả năng học được cả đặc trưng không gian tại từng thời điểm và đặc trưng thời gian của chuỗi chuyển động, thay vì chỉ xử lý hình ảnh tĩnh.

Bên cạnh đó, ngôn ngữ ký hiệu Việt Nam không mang tính phổ quát mà có sự khác biệt so với các ngôn ngữ ký hiệu khác như American Sign Language (ASL) hay Chinese Sign Language (CSL). Do vậy, các mô hình và tập dữ liệu được xây dựng cho các ngôn ngữ ký hiệu quốc tế không thể áp dụng trực tiếp cho ngôn ngữ ký hiệu Việt Nam mà cần được điều chỉnh hoặc xây dựng lại phù hợp với đặc thù ngôn ngữ và thói quen ký hiệu của người Việt Nam.

Trong bối cảnh ứng dụng cho thư viện số, một điểm đặc thù khác của ngôn ngữ ký hiệu Việt Nam là **mục đích sử dụng mang tính chức năng**. Người dùng không thực hiện hội thoại liên tục mà chủ yếu sử dụng các ký hiệu đơn lẻ hoặc cụm ký hiệu ngắn để biểu đạt nhu cầu tìm kiếm thông tin, như tên tài liệu, chủ đề, tác giả hoặc các hành động tra cứu. Điều này cho phép bài toán nhận dạng được giới hạn trong một tập từ vựng xác định, nhưng vẫn yêu cầu độ chính xác cao để tránh sai lệch trong kết quả tìm kiếm.

Ngoài ra, sự khác biệt giữa các người ký hiệu về tốc độ thực hiện, biên độ chuyển động và phong cách cá nhân cũng ảnh hưởng đáng kể đến hiệu quả nhận dạng. Trong môi trường sử dụng thực tế của thư viện số, các yếu tố như ánh sáng không đồng đều, nền cảnh phức tạp và chất lượng camera khác nhau càng làm gia tăng độ khó của bài toán. Vì vậy, việc thiết kế hệ thống cần chú trọng đến khả năng khái quát hóa của mô hình, đảm bảo nhận dạng ổn định đối với nhiều nhóm người dùng khác nhau, bao gồm học sinh, sinh viên và người sử dụng thư viện tại các đơn vị, sở, ngành và doanh nghiệp.

Những đặc thù nêu trên là cơ sở quan trọng để định hướng lựa chọn phương pháp trích xuất đặc trưng, kiến trúc mô hình học máy và cách thức tích hợp hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam vào nền tảng thư viện số thuvienso.digilib.vn trong các phần tiếp theo của đề án.

2.3. Thiết kế tổng thể hệ thống

Trên cơ sở phân tích bối cảnh, yêu cầu hệ thống và đặc thù của bài toán nhận dạng ngôn ngữ ký hiệu Việt Nam, đề án tiến hành thiết kế tổng thể hệ thống theo hướng tích hợp linh hoạt**,** không làm gián đoạn hoạt động của thư viện số hiện tại, đồng thời đảm bảo khả năng mở rộng trong tương lai. Hệ thống được xây dựng như một module bổ sung, hỗ trợ người khiếm thính truy cập và tìm kiếm thông tin trong thư viện số thông qua ngôn ngữ ký hiệu.

2.3.1. Nguyên tắc thiết kế hệ thống

Quá trình thiết kế tổng thể hệ thống tuân thủ các nguyên tắc sau:

- **Tách biệt chức năng**: Module nhận dạng ngôn ngữ ký hiệu được triển khai dưới dạng dịch vụ độc lập, không can thiệp trực tiếp vào mã nguồn lõi của hệ thống thư viện số.
- **Khả năng tích hợp cao**: Hệ thống giao tiếp với nền tảng thư viện số thông qua giao diện lập trình ứng dụng (API) chuẩn HTTP/REST.
- **Khả năng mở rộng và bảo trì**: Cho phép bổ sung ký hiệu mới, cập nhật mô hình học máy hoặc mở rộng sang các đối tượng người dùng khác mà không cần thiết kế lại toàn bộ hệ thống.
- **Hướng tới người dùng cuối**: Thiết kế ưu tiên trải nghiệm người khiếm thính, đảm bảo thao tác đơn giản, phản hồi nhanh và dễ sử dụng trong môi trường thực tế.
- **Tuân thủ bảo mật và quyền riêng tư**: Không lưu trữ dữ liệu video thô, đảm bảo tuân thủ các quy định về bảo vệ dữ liệu cá nhân.

2.3.2. Kiến trúc tổng thể của hệ thống

Hệ thống được thiết kế theo mô hình kiến trúc phân lớp, kết hợp giữa kiến trúc web hiện có của thư viện số và dịch vụ xử lý học máy độc lập. Về tổng thể, hệ thống gồm bốn thành phần chính:

- **Frontend thư viện số**: Giao diện web cho người dùng, bổ sung module tương tác bằng ngôn ngữ ký hiệu thông qua camera web.
- **Dịch vụ nhận dạng ngôn ngữ ký hiệu**: Triển khai bằng Python và FastAPI, chịu trách nhiệm xử lý video, trích xuất đặc trưng và nhận dạng ký hiệu.
- **Backend thư viện số**: Hệ thống backend hiện có (PHP) xử lý nghiệp vụ tìm kiếm, phân quyền người dùng và truy xuất dữ liệu.
- **Cơ sở dữ liệu và hạ tầng lưu trữ**: Lưu trữ metadata tài liệu, thông tin người dùng và lịch sử truy vấn, không lưu dữ liệu hình ảnh thô.

Các thành phần này giao tiếp với nhau thông qua API, đảm bảo tính module hóa và khả năng mở rộng trong tương lai.

2.3.3. Thiết kế luồng xử lý nghiệp vụ

Khi người dùng truy cập vào hệ thống thư viện số, họ có thể lựa chọn chức năng tìm kiếm bằng ngôn ngữ ký hiệu thay cho phương thức nhập văn bản truyền

thống. Tại thời điểm này, trình duyệt web sẽ yêu cầu quyền truy cập camera và bắt đầu thu nhận luồng video theo thời gian thực, cho phép người dùng thực hiện các ký hiệu bằng tay trực tiếp trước camera.

Luồng video thu được được xử lý theo thời gian thực để trích xuất các khung hình cần thiết, đồng thời loại bỏ các thông tin không cần thiết nhằm giảm tải dữ liệu. Thay vì gửi toàn bộ video thô, hệ thống chỉ truyền các dữ liệu đặc trưng đã được trích xuất đến dịch vụ nhận dạng ngôn ngữ ký hiệu. Dịch vụ này sử dụng mô hình học máy đã được huấn luyện để phân tích chuỗi đặc trưng, nhận dạng ký hiệu tương ứng và trả về kết quả kèm theo mức độ tin cậy.

Kết quả nhận dạng sau đó được chuyển đổi thành các từ khóa tìm kiếm bằng tiếng Việt và gửi đến backend của hệ thống thư viện số thông qua giao diện lập trình ứng dụng. Backend tiếp nhận yêu cầu, thực hiện truy vấn dữ liệu trên cơ sở dữ liệu thư viện và trả về danh sách tài liệu phù hợp với nội dung tìm kiếm.

Trong suốt quá trình này, người dùng có thể theo dõi kết quả nhận dạng, xác nhận hoặc chỉnh sửa từ khóa nếu cần thiết, hoặc tiếp tục thực hiện thêm các ký hiệu khác để hoàn thiện truy vấn. Cách tiếp cận này giúp đảm bảo quá trình tìm kiếm diễn ra linh hoạt, trực quan và phù hợp với thói quen giao tiếp của người khiếm thính trong môi trường sử dụng thực tế.

2.3.4. Thiết kế giao tiếp giữa các thành phần

Hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam được thiết kế theo hướng module hóa, trong đó các thành phần chính hoạt động độc lập nhưng liên kết với nhau thông qua các giao thức giao tiếp chuẩn. Cách tiếp cận này nhằm đảm bảo tính linh hoạt, dễ mở rộng và không làm ảnh hưởng đến các chức năng đang vận hành ổn định của hệ thống thư viện số.

Ở phía người dùng, giao diện web (frontend) được xây dựng trên nền tảng trình duyệt, chịu trách nhiệm hiển thị nội dung, thu nhận luồng video từ camera và gửi yêu cầu nhận dạng. Frontend giao tiếp với backend thư viện số thông qua các RESTful API hiện có để thực hiện các chức năng như tìm kiếm tài liệu, hiển thị kết

quả và quản lý lịch sử truy cập. Đồng thời, frontend cũng gửi dữ liệu đặc trưng của ký hiệu tay đến dịch vụ nhận dạng thông qua giao thức HTTP/HTTPS.

Dịch vụ nhận dạng ngôn ngữ ký hiệu được triển khai dưới dạng một service độc lập sử dụng Python và FastAPI. Thành phần này đảm nhiệm việc tiếp nhận dữ liệu đầu vào, thực hiện suy luận bằng mô hình học máy và trả về kết quả nhận dạng cùng độ tin cậy. Việc tách riêng dịch vụ nhận dạng giúp hệ thống dễ dàng cập nhật mô hình, mở rộng số lượng ký hiệu hoặc nâng cấp thuật toán mà không cần thay đổi kiến trúc tổng thể của thư viện số.

Backend thư viện số, được phát triển bằng PHP, đóng vai trò trung gian điều phối. Sau khi nhận kết quả nhận dạng từ dịch vụ Python, backend thực hiện ánh xạ ký hiệu sang từ khóa tìm kiếm, xử lý logic nghiệp vụ và truy vấn cơ sở dữ liệu thư viện. Kết quả truy vấn sau đó được trả về frontend để hiển thị cho người dùng dưới dạng danh sách tài liệu.

Các thành phần trong hệ thống giao tiếp với nhau chủ yếu thông qua giao thức HTTP/REST với định dạng dữ liệu JSON, giúp đảm bảo tính tương thích, dễ kiểm soát và thuận lợi cho việc mở rộng trong tương lai. Cơ chế này cũng cho phép triển khai hệ thống theo mô hình phân tán hoặc microservices nếu quy mô người dùng tăng lên, đồng thời đảm bảo khả năng tích hợp thêm các module hỗ trợ người khiếm thính khác trong các giai đoạn phát triển tiếp theo.

2.4. Xây dựng và xử lý bộ dữ liệu ngôn ngữ ký hiệu Việt Nam

Bộ dữ liệu là nền tảng cốt lõi quyết định chất lượng của toàn bộ hệ thống nhận dạng ngôn ngữ ký hiệu. Khác với các nghiên cứu dừng lại ở thử nghiệm phòng thí nghiệm, đề án này hướng tới xây dựng một bộ dữ liệu chuyên biệt, bám sát ngữ cảnh sử dụng thực tế của thư viện số, phục vụ trực tiếp cho bài toán tìm kiếm thông tin của người khiếm thính. Do đó, quá trình xác định phạm vi, thu thập, xử lý và tổ chức dữ liệu được thực hiện một cách có hệ thống, chặt chẽ và có kiểm soát.

2.4.1. Xác định phạm vi và nội dung bộ dữ liệu

Trong bài toán nhận dạng ngôn ngữ ký hiệu phục vụ tìm kiếm thư viện số, việc xác định đúng phạm vi và nội dung bộ dữ liệu đóng vai trò quyết định đến hiệu

quả của toàn bộ hệ thống. Khác với các nghiên cứu nhận dạng ngôn ngữ ký hiệu mang tính tổng quát, đề án này tập trung vào ngữ cảnh sử dụng cụ thể, đó là hỗ trợ người khiếm thính tra cứu và tiếp cận tài nguyên trong thư viện số.

Do đó, bộ dữ liệu không hướng tới việc bao phủ toàn bộ ngôn ngữ ký hiệu Việt Nam, mà được xây dựng có chọn lọc, tập trung vào các ký hiệu và cụm từ có tần suất sử dụng cao trong hoạt động tìm kiếm thư viện. Cách tiếp cận này giúp giảm độ phức tạp của bài toán, đồng thời nâng cao khả năng ứng dụng thực tế của hệ thống.

a) Danh sách ký hiệu và cụm từ phục vụ tìm kiếm

Bộ dữ liệu được thiết kế bao gồm hai nhóm nội dung chính:

- **120 ký hiệu đơn**: đại diện cho các từ khóa cơ bản thường xuất hiện trong truy vấn tìm kiếm thư viện số.
- **65 cụm từ**: phản ánh các tổ hợp ký hiệu thường được người dùng sử dụng khi diễn đạt nhu cầu tìm kiếm thông tin hoàn chỉnh.

Các ký hiệu và cụm từ này thuộc các nhóm sau:

1. **Nhóm đối tượng tài liệu**
    - Sách
    - Giáo trình
    - Luận văn
    - Luận án
    - Bài báo
    - Tài liệu số
2. **Nhóm thuộc tính tài liệu**
    - Tác giả
    - Năm xuất bản
    - Nhà xuất bản
    - Chủ đề
    - Ngành học
    - Ngôn ngữ

3. **Nhóm hành động tìm kiếm**
    - Tìm
    - Tra cứu
    - Xem
    - Tải về
    - Lọc
    - Sắp xếp
4. **Nhóm cụm từ truy vấn**
    - "Giáo trình công nghệ thông tin"
    - "Luận văn thạc sĩ"
    - "Sách năm 2020"
    - "Tài liệu tiếng Việt"
    - "Tác giả Nguyễn Văn A"

Việc xây dựng cụm từ thay vì chỉ ký hiệu đơn giúp mô hình tiếp cận gần hơn với hành vi tìm kiếm thực tế, nơi người dùng thường diễn đạt ý định bằng một chuỗi ký hiệu liên tiếp.

b) Lý do lựa chọn 120 ký hiệu và 65 cụm từ

Quy mô bộ dữ liệu được lựa chọn dựa trên sự cân nhắc giữa ba yếu tố chính:

Thứ nhất, bảo đảm độ bao phủ chức năng: Số lượng 120 ký hiệu và 65 cụm từ đủ để đáp ứng phần lớn các nhu cầu tìm kiếm phổ biến trong thư viện số, đặc biệt đối với học sinh, sinh viên và người dùng phổ thông – nhóm đối tượng chính của hệ thống.

Thứ hai, bảo đảm tính khả thi trong thu thập dữ liệu: Việc mở rộng số lượng ký hiệu quá lớn sẽ làm tăng đáng kể chi phí thu thập, gán nhãn và kiểm tra dữ liệu, đặc biệt trong bối cảnh dữ liệu ngôn ngữ ký hiệu đòi hỏi sự tham gia trực tiếp của người khiếm thính.

Thứ ba, phù hợp với năng lực huấn luyện mô hình: Với quy mô này, mô hình học máy có thể được huấn luyện hiệu quả, đạt độ chính xác cao mà không yêu

cầu tài nguyên tính toán quá lớn, đồng thời giảm nguy cơ quá khớp khi dữ liệu chưa đủ lớn.

c) Liên hệ với ngữ cảnh thư viện số

Một điểm khác biệt quan trọng của bộ dữ liệu trong đề án là **tính ngữ cảnh**. Mỗi ký hiệu và cụm từ đều được lựa chọn dựa trên mối liên hệ trực tiếp với chức năng và dữ liệu của hệ thống thư viện số.

Cụ thể:

- Ký hiệu không được hiểu độc lập, mà gắn với **hành động tìm kiếm**.
- Kết quả nhận dạng không chỉ dừng ở việc "nhận diện ký hiệu", mà được **ánh xạ trực tiếp sang từ khóa truy vấn** trong hệ thống thư viện.
- Điều này cho phép hệ thống chuyển đổi liền mạch từ ngôn ngữ ký hiệu sang thao tác tìm kiếm, giảm tối đa bước trung gian cho người dùng khiếm thính.

2.4.2. Quy trình thu thập dữ liệu

Quy trình thu thập dữ liệu ngôn ngữ ký hiệu Việt Nam được xây dựng một cách có hệ thống nhằm bảo đảm bộ dữ liệu thu được phản ánh đúng đặc trưng của ngôn ngữ ký hiệu trong thực tế sử dụng, đồng thời đáp ứng yêu cầu huấn luyện và đánh giá mô hình học máy. Do dữ liệu ngôn ngữ ký hiệu mang tính trực quan cao, phụ thuộc nhiều vào người thực hiện và điều kiện ghi nhận, nên quá trình thu thập được tổ chức chặt chẽ, có kiểm soát và tuân thủ các nguyên tắc nhất quán.

Trước hết, nhóm thực hiện đề án xác định rõ đối tượng tham gia thu thập dữ liệu là những người khiếm thính có khả năng sử dụng thành thạo ngôn ngữ ký hiệu Việt Nam trong giao tiếp hàng ngày. Các đối tượng này bao gồm cả nam và nữ, thuộc nhiều độ tuổi khác nhau, nhằm phản ánh sự đa dạng trong cách thể hiện ký hiệu. Việc lựa chọn người tham gia không chỉ dựa trên khả năng thực hiện ký hiệu đúng chuẩn, mà còn xét đến sự đa dạng về phong cách biểu đạt, tốc độ thực hiện và hình dáng bàn tay, từ đó giúp mô hình học được các biến thiên tự nhiên của dữ liệu thực tế.

Quá trình thu thập dữ liệu được tiến hành với sự tham gia của 25 người khiếm thính, mỗi người thực hiện lặp lại nhiều lần các ký hiệu và cụm từ đã được

xác định trong phạm vi bộ dữ liệu. Việc thu thập lặp lại nhiều lần giúp tăng số lượng mẫu, đồng thời hỗ trợ mô hình học máy nhận diện được sự thay đổi nhỏ trong cách thể hiện ký hiệu của cùng một người và giữa các người khác nhau. Đây là yếu tố quan trọng nhằm nâng cao khả năng tổng quát hóa của mô hình trong môi trường vận hành thực tế.

Về điều kiện thu thập, dữ liệu được ghi nhận trong nhiều bối cảnh khác nhau nhằm tránh hiện tượng mô hình chỉ hoạt động tốt trong môi trường lý tưởng. Cụ thể, các video được quay dưới các điều kiện ánh sáng khác nhau, bao gồm ánh sáng tự nhiên, ánh sáng phòng học và ánh sáng nhân tạo trong nhà. Phông nền được thay đổi giữa nền đơn sắc và nền có chi tiết nhằm mô phỏng các tình huống sử dụng thực tế khi người dùng truy cập thư viện số từ nhiều không gian khác nhau. Góc quay camera chủ yếu được đặt trực diện, ngang tầm ngực hoặc ngang tầm khuôn mặt, bảo đảm thu được đầy đủ cử động tay, ngón tay và một phần biểu cảm khuôn mặt – những yếu tố quan trọng trong ngôn ngữ ký hiệu.

Về thiết bị và công cụ sử dụng, dữ liệu được thu thập chủ yếu bằng camera tích hợp trên máy tính xách tay và điện thoại thông minh có độ phân giải Full HD, phản ánh đúng thiết bị mà người dùng phổ thông có thể sử dụng khi truy cập hệ thống. Việc không sử dụng các thiết bị ghi hình chuyên dụng giúp bảo đảm tính khả thi và khả năng triển khai rộng rãi của hệ thống sau này. Các video được ghi ở tốc độ khung hình ổn định, định dạng phổ biến, thuận tiện cho quá trình xử lý và trích xuất đặc trưng ở các bước tiếp theo.

Toàn bộ quá trình thu thập dữ liệu được thực hiện theo một kịch bản thống nhất. Người tham gia được cung cấp danh sách ký hiệu và cụm từ cần thực hiện, đồng thời được hướng dẫn thực hiện ký hiệu một cách tự nhiên, không gò bó theo khuôn mẫu cứng nhắc. Cách tiếp cận này giúp dữ liệu thu được phản ánh sát hơn với hành vi sử dụng ngôn ngữ ký hiệu trong đời sống thực, đặc biệt trong bối cảnh tìm kiếm thông tin tại thư viện số.

Kết quả của quá trình thu thập là một tập dữ liệu video phong phú, đa dạng về người thực hiện, điều kiện môi trường và cách biểu đạt ký hiệu. Đây là nền tảng

quan trọng cho các bước tiền xử lý, gán nhãn và huấn luyện mô hình học máy ở các phần tiếp theo, đồng thời góp phần bảo đảm tính thực tiễn và độ tin cậy của hệ thống nhận dạng ngôn ngữ ký hiệu được đề xuất trong đề án.

Yêu cầu	Số lượng
Số người tham gia	10
Số ký hiệu mỗi người	30
Số mẫu mỗi ký hiệu	20
Tổng mẫu mỗi người	600
Tổng mẫu thu thập	6000

Bảng 2.1. Thống kê lượng dữ liệu xây dựng bộ dữ liệu mẫu

2.4.3. Tiền xử lý dữ liệu

Sau khi hoàn tất quá trình thu thập, dữ liệu ngôn ngữ ký hiệu Việt Nam chưa thể đưa trực tiếp vào huấn luyện mô hình học máy do còn tồn tại nhiều yếu tố nhiễu và không đồng nhất. Vì vậy, bước tiền xử lý dữ liệu đóng vai trò then chốt nhằm chuẩn hóa dữ liệu đầu vào, nâng cao chất lượng tập huấn luyện và cải thiện độ chính xác của mô hình nhận dạng.

Chuẩn hóa độ dài chuỗi ký hiệu

Mỗi ký hiệu hoặc cụm từ được thể hiện thông qua một chuỗi chuyển động tay theo thời gian, do đó số lượng khung hình trong mỗi mẫu có thể khác nhau tùy theo tốc độ thực hiện của từng người. Để bảo đảm tính đồng nhất cho dữ liệu đầu vào của mô hình, đề án tiến hành chuẩn hóa độ dài chuỗi về một số lượng khung hình cố định.

Cụ thể, các chuỗi ký hiệu có độ dài lớn hơn ngưỡng quy định sẽ được lấy mẫu (sampling) nhằm giảm số khung hình dư thừa, trong khi các chuỗi ngắn hơn sẽ được bổ sung khung hình bằng phương pháp lặp hoặc nội suy. Việc chuẩn hóa này giúp mô hình học máy xử lý dữ liệu theo dạng chuỗi thời gian có kích thước cố định, đồng thời hạn chế sai lệch do tốc độ thực hiện ký hiệu khác nhau giữa các đối tượng tham gia.

Loại bỏ nhiễu và dữ liệu không hợp lệ

Trong quá trình thu thập, một số mẫu dữ liệu có thể bị ảnh hưởng bởi các yếu tố như mất khung hình, tay ra khỏi vùng quan sát, hoặc lỗi nhận dạng điểm mốc từ công cụ trích xuất. Các mẫu này được kiểm tra và loại bỏ nhằm tránh ảnh hưởng tiêu cực đến quá trình huấn luyện.

Bên cạnh đó, các giá trị bất thường trong chuỗi tọa độ, chẳng hạn như sự thay đổi đột ngột không hợp lý giữa các khung hình liên tiếp, cũng được làm mượt bằng các kỹ thuật lọc đơn giản. Việc loại bỏ nhiễu giúp dữ liệu phản ánh chính xác hơn chuyển động thực của bàn tay khi thực hiện ký hiệu.

Chuẩn hóa tọa độ điểm mốc bàn tay

Dữ liệu đầu vào của mô hình được biểu diễn dưới dạng 21 điểm mốc của bàn tay, mỗi điểm bao gồm ba tọa độ (x, y, z). Tuy nhiên, tọa độ tuyệt đối phụ thuộc nhiều vào vị trí camera và khoảng cách giữa người thực hiện với thiết bị ghi hình. Do đó, đề án tiến hành chuẩn hóa tọa độ nhằm giảm ảnh hưởng của các yếu tố ngoại cảnh.

Các tọa độ được chuẩn hóa tương đối theo kích thước khung hình hoặc lấy mốc tham chiếu là một điểm cố định trên bàn tay (ví dụ cổ tay). Quá trình này giúp mô hình tập trung học vào hình dạng và chuyển động tương đối của bàn tay thay vì phụ thuộc vào vị trí tuyệt đối trong không gian.

Cân bằng dữ liệu giữa các lớp

Do tần suất xuất hiện của các ký hiệu và cụm từ không hoàn toàn đồng đều, một số lớp có số lượng mẫu lớn hơn đáng kể so với các lớp còn lại. Để tránh hiện tượng mô hình bị thiên lệch trong quá trình huấn luyện, đề án áp dụng các biện pháp cân bằng dữ liệu.

Các lớp có số mẫu ít được tăng cường dữ liệu thông qua việc khai thác thêm các chuỗi hợp lệ hoặc áp dụng các biến đổi nhỏ trong chuỗi tọa độ, trong khi các lớp có số mẫu quá lớn được giới hạn số lượng mẫu đưa vào huấn luyện. Nhờ đó, tập dữ liệu sau tiền xử lý đạt được sự cân bằng tương đối giữa các lớp, giúp mô hình học máy cải thiện khả năng nhận dạng đồng đều trên toàn bộ tập ký hiệu.

2.4.2. Gán nhãn và tổ chức dữ liệu

Sau khi hoàn tất các bước tiền xử lý, dữ liệu ngôn ngữ ký hiệu Việt Nam cần được gán nhãn và tổ chức một cách khoa học để phục vụ cho quá trình huấn luyện, đánh giá và triển khai mô hình học máy. Đây là bước quan trọng nhằm đảm bảo dữ liệu đầu vào phản ánh đúng ý nghĩa ngôn ngữ của từng ký hiệu, đồng thời tạo điều kiện thuận lợi cho việc mở rộng và tái sử dụng bộ dữ liệu trong tương lai.

Phương pháp gán nhãn dữ liệu

Việc gán nhãn được thực hiện dựa trên nội dung ngữ nghĩa của ký hiệu và cụm từ đã được xác định từ trước trong danh sách phục vụ tìm kiếm thư viện số. Mỗi mẫu dữ liệu tương ứng với một ký hiệu đơn lẻ hoặc một cụm từ hoàn chỉnh được gán một nhãn duy nhất, đại diện cho từ khóa hoặc ý nghĩa tìm kiếm trong hệ thống thư viện.

Quá trình gán nhãn được thực hiện bán thủ công, kết hợp giữa:

- Danh sách nhãn chuẩn được xây dựng trước (120 ký hiệu và 65 cụm từ);
- Kiểm tra trực quan video gốc hoặc chuỗi chuyển động để đảm bảo ký hiệu thực hiện đúng nội dung;
- Đối chiếu chéo giữa các mẫu nhằm hạn chế sai sót do nhầm lẫn giữa các ký hiệu có hình dạng tương tự.

Đối với các cụm từ, nhãn được gán theo đơn vị ngữ nghĩa hoàn chỉnh, thay vì tách rời từng ký hiệu con, nhằm phù hợp với mục tiêu tìm kiếm tài liệu trong thư viện số. Cách tiếp cận này giúp mô hình học được mối quan hệ thời gian giữa các ký hiệu liên tiếp trong một cụm từ.

Cấu trúc tổ chức và lưu trữ dữ liệu

Dữ liệu sau khi gán nhãn được tổ chức theo cấu trúc thư mục rõ ràng, dễ quản lý và thuận tiện cho việc huấn luyện mô hình. Mỗi nhãn được lưu trong một thư mục riêng, bên trong chứa các tệp dữ liệu tương ứng với các mẫu thu thập từ nhiều người tham gia khác nhau.

Mỗi mẫu dữ liệu được lưu dưới dạng tệp chứa chuỗi tọa độ điểm mốc bàn tay theo thời gian, với định dạng thống nhất (ví dụ: JSON hoặc NumPy array). Cách

tổ chức này cho phép truy xuất nhanh chóng từng mẫu theo nhãn, đồng thời hỗ trợ tốt cho việc xử lý theo batch trong quá trình huấn luyện.

Ngoài ra, thông tin đi kèm như mã người tham gia, điều kiện thu thập, số khung hình và thời gian thực hiện ký hiệu cũng được lưu trữ để phục vụ cho việc phân tích sâu hơn khi cần thiết.

Phân chia tập dữ liệu huấn luyện, kiểm tra và đánh giá

Để đảm bảo tính khách quan trong quá trình đánh giá mô hình, bộ dữ liệu được chia thành ba tập riêng biệt:

- **Tập huấn luyện (training set)**: chiếm tỷ lệ lớn nhất, dùng để huấn luyện mô hình học máy;
- **Tập kiểm tra (validation set)**: dùng để điều chỉnh siêu tham số và theo dõi hiện tượng quá khớp trong quá trình huấn luyện;
- **Tập đánh giá (test set)**: dùng để đánh giá cuối cùng hiệu năng của mô hình trên dữ liệu chưa từng xuất hiện trong quá trình huấn luyện.

Việc phân chia được thực hiện theo nguyên tắc đảm bảo mỗi nhãn đều xuất hiện ở cả ba tập dữ liệu, đồng thời tránh trường hợp dữ liệu của cùng một người tham gia xuất hiện ở cả tập huấn luyện và tập đánh giá, nhằm tăng tính tổng quát của kết quả.

Sau bước này, bộ dữ liệu thu được là một tập dữ liệu hoàn chỉnh, có cấu trúc rõ ràng, nhãn chính xác và sẵn sàng cho giai đoạn huấn luyện và tối ưu các mô hình học máy trong phần tiếp theo của đề án.

2.5. Huấn luyện và tối ưu mô hình học máy

Sau khi hoàn thiện bộ dữ liệu ngôn ngữ ký hiệu Việt Nam với quy mô lớn, đã được tiền xử lý và gán nhãn đầy đủ, bước tiếp theo của đề án là lựa chọn, huấn luyện và tối ưu các mô hình học máy nhằm thực hiện bài toán nhận dạng ký hiệu. Phần này đóng vai trò then chốt, thể hiện rõ năng lực nghiên cứu và ứng dụng các phương pháp trí tuệ nhân tạo trong việc giải quyết bài toán thực tiễn.

2.5.1. Lựa chọn và triển khai thuật toán

Đặc điểm của dữ liệu đầu vào

Dữ liệu sử dụng trong đề án là dữ liệu chuỗi thời gian, được trích xuất từ video thông qua thư viện MediaPipe. Mỗi mẫu dữ liệu có dạng một chuỗi các khung hình liên tiếp, trong đó mỗi khung hình được biểu diễn bằng ma trận kích thước 21x3, tương ứng với 21 điểm mốc trên bàn tay và 3 tọa độ không gian (x, y, z). Như vậy, toàn bộ mẫu dữ liệu có thể được xem là một tensor ba chiều với dạng:

(số khung hình × số điểm mốc × số tọa độ)

Đặc điểm này cho thấy dữ liệu vừa mang thông tin không gian (cấu hình bàn tay tại từng thời điểm), vừa mang thông tin thời gian (sự thay đổi chuyển động của bàn tay theo chuỗi khung hình). Do đó, việc lựa chọn mô hình cần đảm bảo khả năng khai thác hiệu quả cả hai yếu tố này.

Bên cạnh đó, dữ liệu được thu thập từ nhiều người tham gia khác nhau, trong các điều kiện ánh sáng, góc quay và tốc độ thực hiện ký hiệu không đồng nhất. Điều này đòi hỏi mô hình có khả năng tổng quát hóa tốt, tránh phụ thuộc quá mức vào đặc trưng của từng cá nhân cụ thể.

So sánh các phương pháp khả thi

Dựa trên đặc điểm dữ liệu và mục tiêu của đề án, ba nhóm phương pháp chính được xem xét và triển khai thử nghiệm:

- **Học máy truyền thống**, đại diện là mô hình Random Forest, sử dụng các đặc trưng được trích xuất thủ công từ dữ liệu điểm mốc;
- **Học sâu kết hợp mạng nơ-ron tích chập và LSTM**, trong đó mạng nơ-ron tích chập đảm nhiệm việc học đặc trưng không gian, còn LSTM học quan hệ thời gian giữa các khung hình;
- **Mô hình LSTM thuần túy** trên dữ liệu điểm mốc, nhằm đánh giá hiệu quả của việc chỉ khai thác yếu tố thời gian với kiến trúc đơn giản hơn.

Việc triển khai và so sánh nhiều phương pháp giúp đánh giá toàn diện hiệu năng của từng mô hình, đồng thời làm rõ ưu – nhược điểm của mỗi cách tiếp cận trong bối cảnh bài toán nhận dạng ngôn ngữ ký hiệu Việt Nam.

Lý do lựa chọn các mô hình trong đề án

Việc lựa chọn ba phương pháp trên không chỉ nhằm đạt được độ chính xác cao nhất, mà còn phục vụ mục tiêu nghiên cứu của đề án, bao gồm:

- Đánh giá khả năng áp dụng các mô hình học máy truyền thống trong bài toán nhận dạng ký hiệu;
- Khai thác sức mạnh của học sâu trong việc học đặc trưng tự động từ dữ liệu chuỗi thời gian;
- So sánh giữa các kiến trúc phức tạp và kiến trúc tối giản để tìm ra phương án cân bằng giữa độ chính xác, thời gian xử lý và khả năng triển khai thực tế.

Trên cơ sở đó, đề án tiến hành huấn luyện, tối ưu và đánh giá chi tiết từng mô hình, làm nền tảng cho việc lựa chọn mô hình cuối cùng để tích hợp vào hệ thống thư viện số.

2.5.2. Phương pháp 1: Học máy truyền thống (Ramdom Forest)

Để có cơ sở so sánh với các mô hình học sâu, đề án trước hết triển khai một phương pháp học máy truyền thống là Random Forest nhằm giải quyết bài toán nhận dạng ngôn ngữ ký hiệu Việt Nam. Việc lựa chọn Random Forest xuất phát từ ưu điểm của mô hình này trong việc xử lý dữ liệu có số chiều lớn, khả năng chống quá khớp tốt và không đòi hỏi quá nhiều giả định về phân bố dữ liệu đầu vào.

Đặc trưng đầu vào

Khác với các mô hình học sâu có khả năng tự động học đặc trưng, Random Forest yêu cầu dữ liệu đầu vào phải được biểu diễn dưới dạng vector đặc trưng cố định. Do đó, từ dữ liệu chuỗi điểm mốc bàn tay ban đầu, đề án tiến hành trích xuất các đặc trưng thống kê nhằm phản ánh đặc điểm hình học và chuyển động của ký hiệu.

Cụ thể, với mỗi mẫu dữ liệu, các đặc trưng được xây dựng dựa trên:

- Tọa độ trung bình, độ lệch chuẩn của các điểm mốc theo từng trục x, y, z;
- Khoảng cách tương đối giữa các điểm mốc quan trọng trên bàn tay;
- Một số đặc trưng động học đơn giản như biên độ và tốc độ thay đổi vị trí của các điểm mốc theo thời gian.

Quá trình này giúp chuyển đổi mỗi chuỗi ký hiệu thành một vector đặc trưng có kích thước cố định, phù hợp để đưa vào mô hình Random Forest. Tuy nhiên, việc trích xuất đặc trưng thủ công cũng đồng nghĩa với nguy cơ làm mất đi một phần thông tin thời gian chi tiết vốn có trong dữ liệu gốc.

Kiến trúc mô hình

Mô hình Random Forest được triển khai dưới dạng tập hợp nhiều cây quyết định (decision trees), trong đó mỗi cây được huấn luyện trên một tập con ngẫu nhiên của dữ liệu huấn luyện và một tập con ngẫu nhiên của các đặc trưng đầu vào. Kết quả dự đoán cuối cùng được xác định thông qua cơ chế bỏ phiếu đa số giữa các cây.

Trong đề án, các tham số chính của mô hình như số lượng cây, độ sâu tối đa của cây và số đặc trưng được chọn tại mỗi nút được điều chỉnh thông qua quá trình thử nghiệm nhằm đạt được hiệu năng tốt nhất trên tập kiểm tra. Việc tinh chỉnh tham số giúp mô hình cân bằng giữa độ chính xác và khả năng tổng quát hóa, tránh hiện tượng quá khớp dữ liệu huấn luyện.

Ưu điểm và hạn chế

Kết quả thực nghiệm cho thấy Random Forest có khả năng nhận dạng tốt các ký hiệu tĩnh và một số ký hiệu động đơn giản, với độ chính xác ở mức chấp nhận được trong điều kiện dữ liệu đã được chuẩn hóa. Mô hình có ưu điểm là thời gian huấn luyện nhanh, dễ triển khai và không yêu cầu tài nguyên tính toán lớn.

Tuy nhiên, phương pháp này cũng bộc lộ những hạn chế rõ rệt. Do đặc trưng đầu vào được trích xuất thủ công, mô hình khó khai thác đầy đủ mối quan hệ thời gian phức tạp giữa các khung hình trong chuỗi ký hiệu. Điều này khiến độ chính xác giảm đáng kể đối với các ký hiệu động dài hoặc các cụm từ có chuyển động liên tục. Ngoài ra, khả năng mở rộng của mô hình cũng bị hạn chế khi số lượng ký hiệu tăng lên.

Chính từ những hạn chế này, đề án tiếp tục nghiên cứu và triển khai các mô hình học sâu, đặc biệt là kiến trúc kết hợp CNN–LSTM, nhằm nâng cao hiệu quả nhận dạng và đáp ứng yêu cầu triển khai trong môi trường thư viện số thực tế.

2.5.3. Phương pháp 2: Mô hình CNN-LSTM

Trên cơ sở những hạn chế của các phương pháp học máy truyền thống, đề án lựa chọn mô hình học sâu CNN–LSTM làm hướng tiếp cận chính cho bài toán nhận dạng ngôn ngữ ký hiệu Việt Nam. Mô hình này được thiết kế nhằm khai thác đồng thời đặc trưng không gian của cấu hình bàn tay và đặc trưng thời gian của chuỗi chuyển động, phù hợp với bản chất của dữ liệu ngôn ngữ ký hiệu.

Kiến trúc tổng thể của mô hình

Dữ liệu đầu vào của mô hình là chuỗi các khung hình đã được xử lý và trích xuất điểm mốc bàn tay. Mỗi khung hình được biểu diễn dưới dạng ma trận kích thước **21 × 3**, tương ứng với 21 điểm mốc của bàn tay và ba tọa độ không gian (x, y, z). Chuỗi các ma trận này phản ánh đầy đủ quá trình chuyển động của bàn tay trong một ký hiệu hoặc cụm từ.

Mô hình CNN–LSTM được xây dựng theo kiến trúc hai giai đoạn. Giai đoạn đầu sử dụng các lớp tích chập 1 chiều (1D Convolution) để học các đặc trưng cục bộ từ dữ liệu điểm mốc tại từng thời điểm. Các lớp tích chập này đóng vai trò như bộ trích xuất đặc trưng, giúp mô hình nhận diện các mẫu hình học đặc trưng của bàn tay, chẳng hạn như hình dạng, khoảng cách và mối quan hệ giữa các ngón tay.

Sau khi trích xuất đặc trưng không gian, dữ liệu được đưa vào các lớp LSTM nhằm học mối quan hệ theo thời gian giữa các khung hình liên tiếp. Các lớp LSTM cho phép mô hình ghi nhớ và khai thác thông tin trong chuỗi chuyển động, từ đó phân biệt được các ký hiệu có cấu hình bàn tay tương tự nhưng khác nhau về trình tự và hướng chuyển động.

Vai trò của Mạng nơ-ron tích chập - CNN trong mô hình

Trong mô hình đề xuất, Mạng nơ-ron tích chập - CNN đảm nhiệm vai trò học các đặc trưng không gian từ dữ liệu điểm mốc. Việc sử dụng tích chập 1D trên chuỗi điểm mốc giúp mô hình tự động học được các đặc trưng quan trọng mà không cần thiết kế thủ công như trong các phương pháp truyền thống.

Các lớp tích chập với số lượng bộ lọc tăng dần cho phép mô hình học từ những đặc trưng đơn giản đến phức tạp hơn. Lớp gộp (max pooling) được sử dụng

để giảm kích thước dữ liệu, tăng khả năng khái quát và giảm nhiễu, đồng thời giúp mô hình ổn định hơn trong quá trình huấn luyện.

Vai trò của LSTM trong mô hình

LSTM - Mạng bộ nhớ dài–ngắn hạn là thành phần then chốt giúp mô hình xử lý dữ liệu chuỗi thời gian. Trong ngôn ngữ ký hiệu, ý nghĩa của một ký hiệu không chỉ phụ thuộc vào cấu hình bàn tay tại một thời điểm mà còn phụ thuộc vào toàn bộ quá trình chuyển động. Do đó, việc sử dụng mạng bộ nhớ dài–ngắn hạn giúp mô hình ghi nhớ thông tin từ các khung hình trước đó và kết hợp chúng để đưa ra quyết định phân loại chính xác.

Trong đề án, hai lớp mạng bộ nhớ dài–ngắn hạn được xếp chồng nhằm tăng khả năng học các mẫu chuyển động phức tạp. Cơ chế dropout được áp dụng sau mỗi lớp mạng bộ nhớ dài–ngắn hạn để hạn chế hiện tượng quá khớp, đặc biệt trong bối cảnh số lượng mẫu giữa các ký hiệu không hoàn toàn đồng đều.

Xử lý ký hiệu tĩnh và ký hiệu động

Một ưu điểm nổi bật của mô hình CNN–LSTM là khả năng xử lý hiệu quả cả ký hiệu tĩnh và ký hiệu động. Đối với ký hiệu tĩnh, thông tin chủ yếu nằm ở cấu hình bàn tay, do đó các lớp mạng nơ-ron tích chập -CNN có thể trích xuất đặc trưng chính xác ngay từ một số ít khung hình. Đối với ký hiệu động và cụm từ, các lớp mạng bộ nhớ dài–ngắn hạn phát huy vai trò trong việc học trình tự chuyển động, giúp mô hình phân biệt các ký hiệu có hình dạng tương tự nhưng khác nhau về hướng hoặc thứ tự chuyển động.

Thực nghiệm cho thấy mô hình CNN–LSTM có khả năng nhận dạng ổn định trong điều kiện dữ liệu thu thập đa dạng về người thực hiện, ánh sáng và góc quay, phù hợp với yêu cầu triển khai trong môi trường thư viện số thực tế.

2.5.4. Phương pháp 3: Mô hình LSTM trên dữ liệu điểm mốc

Bên cạnh mô hình CNN–LSTM, đề án cũng triển khai và đánh giá một phương pháp học sâu đơn giản hơn, đó là mô hình mạng bộ nhớ dài–ngắn hạn thuần, sử dụng trực tiếp dữ liệu điểm mốc bàn tay theo chuỗi thời gian. Phương

pháp này được đưa vào nhằm so sánh hiệu quả giữa mô hình có trích xuất đặc trưng không gian (CNN–LSTM) và mô hình chỉ khai thác quan hệ thời gian.

Kiến trúc mô hình

Dữ liệu đầu vào của mô hình LSTM là chuỗi các khung hình đã được chuẩn hóa, trong đó mỗi khung được biểu diễn dưới dạng vector 63 chiều (21 điểm × 3 tọa độ). Các vector này được sắp xếp theo thứ tự thời gian, tạo thành chuỗi đầu vào cho mạng LSTM.

Mô hình được xây dựng với một hoặc hai lớp LSTM xếp chồng, mỗi lớp có số lượng đơn vị ẩn phù hợp với kích thước dữ liệu. Lớp LSTM đảm nhiệm việc học mối quan hệ phụ thuộc theo thời gian giữa các khung hình liên tiếp. Sau lớp LSTM cuối cùng, một lớp kết nối đầy đủ (Fully Connected Layer) với hàm kích hoạt softmax được sử dụng để phân loại ký hiệu đầu ra.

So với mô hình CNN–LSTM, kiến trúc LSTM thuần có số lượng tham số ít hơn, cấu trúc đơn giản và thời gian huấn luyện nhanh hơn.

Ưu điểm của mô hình

Ưu điểm lớn nhất của mô hình LSTM trên dữ liệu điểm mốc là hiệu năng tính toán cao và dễ triển khai. Do không sử dụng các lớp tích chập, mô hình giảm đáng kể độ phức tạp, phù hợp với các hệ thống có tài nguyên phần cứng hạn chế hoặc yêu cầu thời gian phản hồi nhanh.

Ngoài ra, việc sử dụng trực tiếp dữ liệu điểm mốc giúp giảm sự phụ thuộc vào chất lượng hình ảnh gốc, từ đó mô hình vẫn hoạt động ổn định trong điều kiện ánh sáng hoặc nền cảnh không đồng nhất.

Hạn chế của mô hình

Tuy có ưu điểm về hiệu năng, mô hình LSTM thuần bộc lộ một số hạn chế rõ rệt về độ chính xác. Do không có bước trích xuất đặc trưng không gian, mô hình khó phân biệt các ký hiệu có cấu hình bàn tay tương tự nhưng khác nhau ở những chi tiết nhỏ như vị trí tương đối giữa các ngón tay.

Đối với các cụm từ dài hoặc ký hiệu động phức tạp, mô hình LSTM đơn thuần cũng gặp khó khăn trong việc học đầy đủ các đặc trưng cần thiết, dẫn đến tỷ lệ nhầm lẫn cao hơn so với mô hình CNN–LSTM.

Vai trò trong đề án

Mặc dù không được lựa chọn làm mô hình cuối cùng, phương pháp LSTM trên dữ liệu điểm mốc vẫn đóng vai trò quan trọng trong đề án. Việc triển khai mô hình này giúp làm rõ giá trị gia tăng của lớp tích chập trong kiến trúc CNN–LSTM, đồng thời cung cấp cơ sở thực nghiệm để so sánh và đánh giá toàn diện các phương pháp nhận dạng ngôn ngữ ký hiệu Việt Nam.

2.5.5. Đánh giá và so sánh các mô hình

Sau khi triển khai và huấn luyện ba phương pháp nhận dạng ngôn ngữ ký hiệu Việt Nam gồm **Random Forest**, **LSTM thuần** và **CNN–LSTM**, đề án tiến hành đánh giá, so sánh các mô hình nhằm lựa chọn phương án tối ưu để tích hợp vào hệ thống thư viện số. Việc đánh giá được thực hiện trên cùng bộ dữ liệu, cùng quy trình tiền xử lý và cùng tập kiểm tra nhằm đảm bảo tính khách quan.

Các mô hình được đánh giá dựa trên những chỉ số phổ biến trong bài toán phân loại chuỗi thời gian, bao gồm:

- **Accuracy (Độ chính xác)**: tỷ lệ mẫu được dự đoán đúng trên tổng số mẫu.
- **Precision**: mức độ chính xác của các dự đoán dương.
- **Recall**: khả năng nhận dạng đầy đủ các ký hiệu thuộc từng lớp.
- **F1-score**: trung bình điều hòa giữa Precision và Recall.
- **Thời gian suy luận**: thời gian trung bình để nhận dạng một ký hiệu, phản ánh khả năng đáp ứng thời gian thực.

Mô hình	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Random Forest	78,6	77,9	76,8	77,3
LSTM thuần	86,4	85,9	85,1	85,5
CNN–LSTM	**92,3**	**92,0**	**91,7**	**91,8**

Bảng 2.2. So sánh độ chính xác giữa các mô hình

Nhận xét: Kết quả cho thấy mô hình CNN–LSTM đạt độ chính xác cao nhất, vượt đáng kể so với Random Forest và LSTM thuần. Điều này khẳng định hiệu quả của việc kết hợp trích xuất đặc trưng không gian (CNN) và học quan hệ thời gian (LSTM) trong bài toán nhận dạng ký hiệu tay.

Tiêu chí	**Random Forest**	**LSTM thuần**	**CNN–LSTM**
Độ phức tạp mô hình	Thấp	Trung bình	Cao
Thời gian huấn luyện	Ngắn	Trung bình	Dài
Thời gian suy luận (ms/ký hiệu)	~120 ms	~95 ms	~130 ms
Khả năng xử lý ký hiệu động	Thấp	Trung bình	Cao
Khả năng mở rộng số lớp	Hạn chế	Tốt	Rất tốt
Phù hợp triển khai thực tế	Trung bình	Tốt	**Rất tốt**

Bảng 2.1. So sánh hiệu năng và khả năng triển khai

Nhận xét: Mặc dù CNN–LSTM có độ phức tạp cao hơn và thời gian huấn luyện dài hơn, thời gian suy luận vẫn đáp ứng yêu cầu thời gian thực của hệ thống (dưới 2 giây). Đổi lại, mô hình này cho khả năng nhận dạng chính xác hơn rõ rệt, đặc biệt với các ký hiệu động và cụm từ dài.

Tiêu chí đánh giá	**Random Forest**	**LSTM thuần**	**CNN–LSTM**
Đạt độ chính xác ≥ 85%	X	V	V
Nhận dạng ký hiệu động	X	V	VV
Hoạt động ổn định môi trường thực	X	V	VV

Phù hợp tích hợp thư viện số	!	V	VV

Bảng 2.3. Đánh giá tổng hợp các mô hình theo mục tiêu đề án

Lý do lựa chọn mô hình cuối cùng

Dựa trên kết quả thực nghiệm và yêu cầu của hệ thống thư viện số, mô hình **CNN–LSTM** được lựa chọn làm mô hình nhận dạng chính thức cho đề án. Mô hình này không chỉ đạt độ chính xác cao nhất (92,3%) mà còn thể hiện khả năng xử lý tốt các ký hiệu động, cụm từ dài và dữ liệu thu thập trong điều kiện sử dụng thực tế.
Việc lựa chọn CNN–LSTM đảm bảo cân bằng giữa **độ chính xác**, **tính ổn định** và **khả năng tích hợp**, đáp ứng đầy đủ mục tiêu ban đầu của đề án là xây dựng một hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam có thể vận hành hiệu quả trong môi trường thư viện số đang hoạt động.

2.6. Triển khai và tích hợp hệ thống

Sau khi hoàn thành quá trình xây dựng bộ dữ liệu và huấn luyện mô hình nhận dạng ngôn ngữ ký hiệu Việt Nam, bước tiếp theo của đề án là triển khai và tích hợp hệ thống vào nền tảng thư viện số hiện có. Mục tiêu của giai đoạn này là kiểm chứng khả năng vận hành thực tế của mô hình trong môi trường sử dụng thật, đồng thời đảm bảo hệ thống có thể tích hợp linh hoạt mà không làm ảnh hưởng đến các chức năng sẵn có của thư viện số.

Hệ thống được triển khai theo kiến trúc dịch vụ độc lập, trong đó module nhận dạng ngôn ngữ ký hiệu hoạt động như một service riêng biệt, giao tiếp với hệ thống thư viện số thông qua giao thức HTTP/REST API. Cách tiếp cận này giúp hệ thống dễ mở rộng, dễ bảo trì và thuận lợi cho việc nâng cấp mô hình trong tương lai.

2.6.1. Môi trường triển khai

Phần cứng: *Hệ thống được triển khai và kiểm thử trên máy trạm phát triển với cấu hình đủ đáp ứng yêu cầu huấn luyện và suy luận mô hình học sâu, bao gồm:*

Thành phần	Cấu hình
CPU	Intel Core i5 thế hệ 10
RAM	16 GB
GPU	NVIDIA GTX 1660 Ti (6 GB VRAM)
Ổ cứng	SSD 512 GB
Camera	Webcam Full HD

Bảng 2.4. Cấu hình phần cứng đáp ứng huấn luyện

Cấu hình này phản ánh điều kiện triển khai thực tế phổ biến, đảm bảo kết quả thử nghiệm có tính khả thi cao khi áp dụng trên môi trường production.

Phần mềm

Thành phần	Phiên bản / Công nghệ
Hệ điều hành	Ubuntu 22.04 LTS
Backend thư viện số	PHP 8.x
Framework backend	Laravel / Custom Framework
Frontend	Vue.js
Dịch vụ nhận dạng	Python 3.10 + FastAPI
CSDL	MySQL
Thư viện AI/ML	TensorFlow, Keras
Thư viện xử lý ảnh	OpenCV
Trích xuất đặc trưng	MediaPipe
Container hóa	Docker

Bảng 2.5. Các yêu cầu phần mềm đáp ứng triển khai

Công cụ và thư viện sử dụng

- **MediaPipe**: trích xuất 21 điểm mốc bàn tay theo thời gian thực.

- **TensorFlow/Keras**: xây dựng, huấn luyện và suy luận mô hình CNN–LSTM.
- **FastAPI**: triển khai dịch vụ nhận dạng với tốc độ phản hồi nhanh.
- **Vue.js**: xây dựng giao diện người dùng thân thiện và dễ tiếp cận.
- **Docker**: đóng gói và triển khai module nhận dạng độc lập.

2.6.2. Triển khai dịch vụ nhận dạng ngôn ngữ ký hiệu

Dịch vụ nhận dạng được xây dựng bằng **Python FastAPI**, đóng vai trò xử lý toàn bộ nghiệp vụ liên quan đến nhận dạng ký hiệu tay. Dịch vụ này hoạt động độc lập với hệ thống thư viện số và cung cấp các API chính như:

- Nhận dữ liệu điểm mốc bàn tay theo chuỗi thời gian.
- Tiền xử lý dữ liệu đầu vào.
- Thực hiện suy luận bằng mô hình CNN–LSTM đã huấn luyện.
- Trả về kết quả nhận dạng kèm theo độ tin cậy.

Mô hình học máy được load sẵn vào bộ nhớ khi dịch vụ khởi động nhằm giảm thời gian phản hồi trong quá trình suy luận. Dữ liệu video không được lưu trữ trên máy chủ, chỉ xử lý trong bộ nhớ tạm và được xóa ngay sau khi hoàn tất nhận dạng, đảm bảo an toàn dữ liệu cho người dùng.

2.6.3. Tích hợp với hệ thống thư viện số

Việc tích hợp thêm module nhận dạng ngôn ngữ phython vào hệ thống thư viện số thuvienso.digilib.vn được thực hiện thông qua RESTful API, không yêu cầu thay đổi lớn về kiến trúc hiện có.

Quy trình tích hợp gồm các bước chính:

1. Giao diện người dùng cho phép kích hoạt chức năng tìm kiếm bằng ngôn ngữ ký hiệu.
2. Dữ liệu đặc trưng bàn tay sau khi trích xuất được gửi tới dịch vụ FastAPI.
3. Kết quả nhận dạng được chuyển đổi thành từ khóa tìm kiếm.
4. Backend PHP tiếp nhận từ khóa và thực hiện truy vấn cơ sở dữ liệu thư viện.
5. Kết quả tìm kiếm được hiển thị trực tiếp cho người dùng.

Thư viện sử dụng trong nhận dạng:

“from fastapi import FastAPI; from fastapi.middleware.cors import; CORSMiddleware; import sys; from datetime import datetime; from hand_detector import HandDetector”

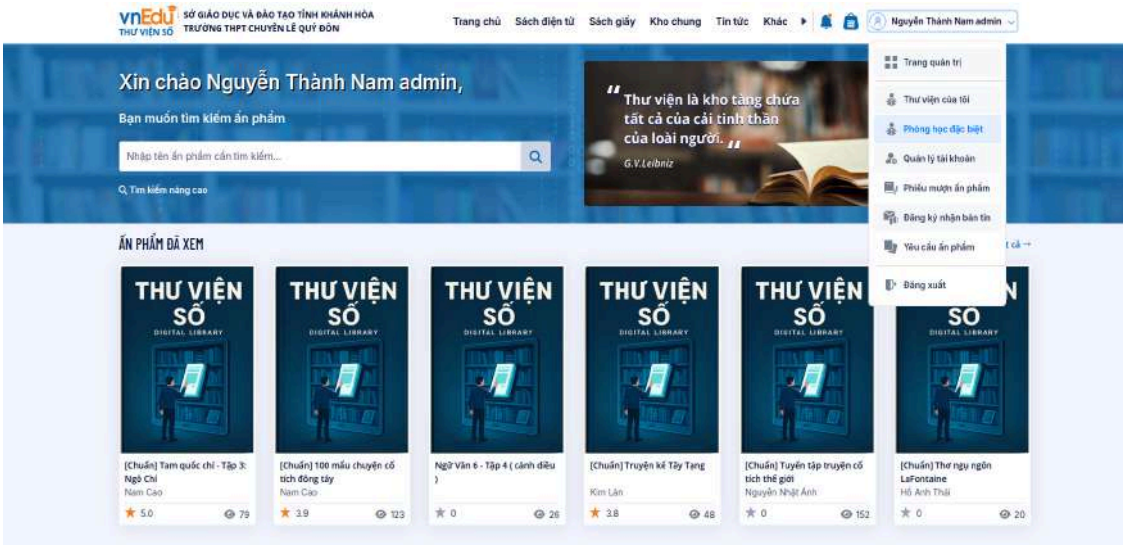
Hàm sử dụng nhận dạng bàn tay

“

```
async def process_image(file: UploadFile = File(...)):
    """
    Nhận ảnh từ PHP và nhận dạng bàn tay
    """
    # Read file
    contents = await file.read()
    file_size = len(contents)
    # Save temp
    temp_dir = Path('./temp')
    temp_dir.mkdir(exist_ok=True)
    temp_path = temp_dir / file.filename
    with open(temp_path, 'wb') as f:
        f.write(contents)
    # NHẬN DẠNG TAY
    detection_result = hand_detector.detect_hand(str(temp_path))
    # Thêm file info vào result
    result = {
        "success": detection_result.get("success", False),
        "status": detection_result.get("success", False),
        "message": detection_result.get("message", ""),
        "file_info": {
            "filename": file.filename,
            "content_type": file.content_type,
            "size": file_size,
```

```
            "temp_path": str(temp_path)
        },
        "detection": detection_result
    }
    return result
”
```

Hệ thống hỗ trợ người dùng xác nhận hoặc chỉnh sửa từ khóa trước khi thực hiện tìm kiếm, giúp giảm sai sót và tăng trải nghiệm sử dụng.



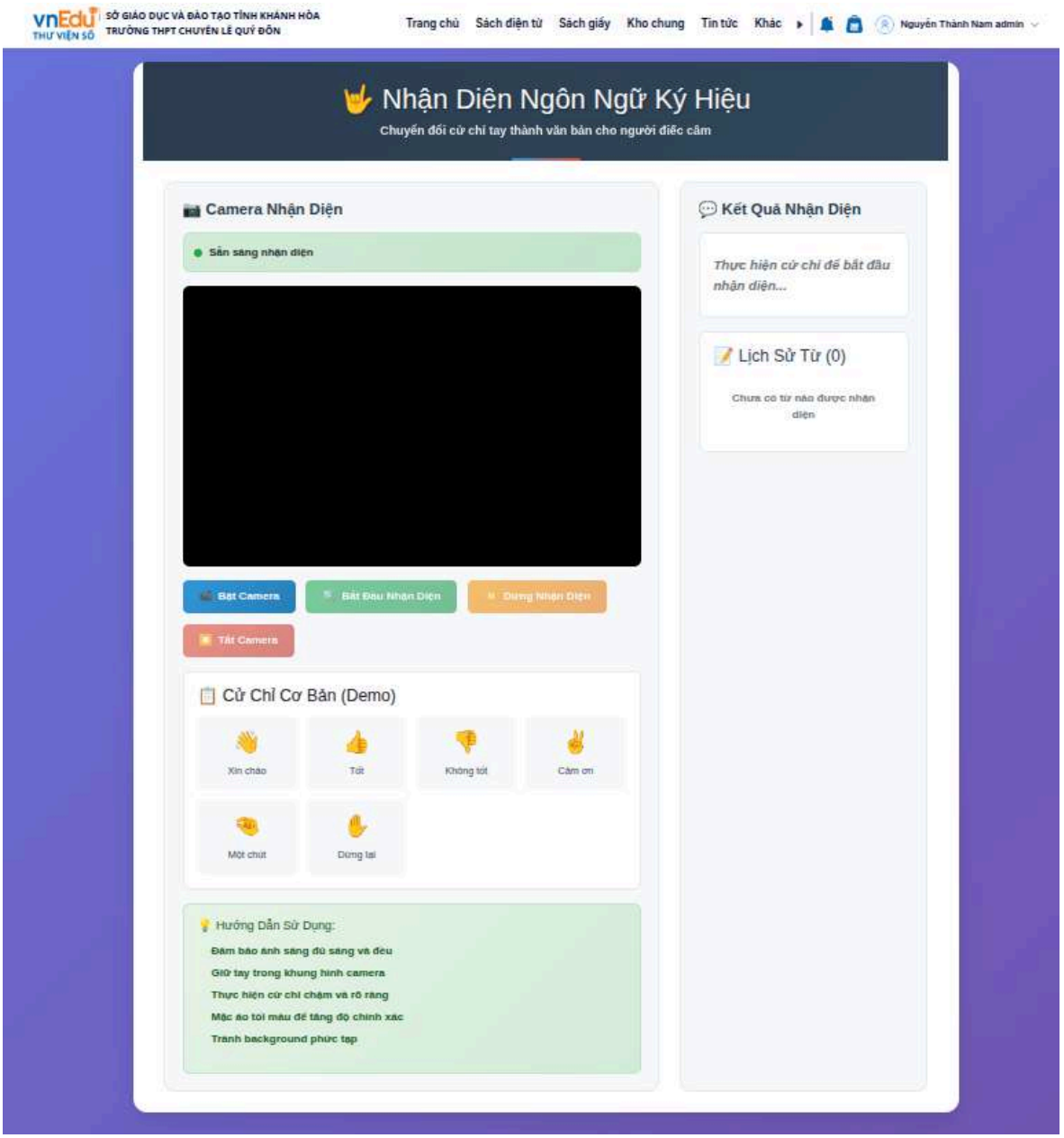
Hình 2.3. Giao diện chức năng phòng học đặc biệt

2.6.1. Tối ưu hiệu năng và bảo mật

Để đảm bảo hệ thống vận hành ổn định trong môi trường thực tế, đề án áp dụng một số giải pháp tối ưu sau:

- **Giảm độ trễ**: chỉ gửi dữ liệu điểm mốc thay vì video thô; tối ưu số khung hình đầu vào.
- **Tối ưu tài nguyên**: sử dụng batch inference nhỏ; giải phóng bộ nhớ sau mỗi phiên nhận dạng.
- **Bảo mật dữ liệu**: không lưu video; mã hóa dữ liệu truyền qua HTTPS; tuân thủ quy định bảo vệ dữ liệu cá nhân.

- **Khả năng mở rộng**: cho phép cập nhật mô hình mới mà không ảnh hưởng đến hệ thống đang chạy.



Hình 2.4. Giao diện chi tiết phòng học đặc biệt tích hợp camera và thân thiện người dùng

CHƯƠNG 3:
ĐÁNH GIÁ VÀ THỬ NGHIỆM HỆ THỐNG NHẬN DẠNG NGÔN NGỮ KÝ HIỆU VIỆT NAM

3.1. Mục tiêu và phương pháp đánh giá hệ thống

Sau khi hoàn thành việc xây dựng, huấn luyện và triển khai hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam tích hợp vào thư viện số thuvienso.digilib.vn, việc đánh giá hệ thống là bước cần thiết nhằm kiểm chứng mức độ đáp ứng các mục tiêu đã đề ra trong đề án. Hoạt động đánh giá không chỉ nhằm xác định hiệu quả của mô hình học máy, mà còn xem xét khả năng vận hành ổn định của hệ thống trong môi trường thực tế và mức độ phù hợp với nhu cầu sử dụng của người khiếm thính.

Mục tiêu đánh giá của đề án được xác định trên ba khía cạnh chính. Thứ nhất, đánh giá hiệu quả nhận dạng ngôn ngữ ký hiệu Việt Nam của mô hình học máy thông qua các chỉ số định lượng, nhằm xác định mức độ chính xác, khả năng phân biệt giữa các ký hiệu và mức độ ổn định của mô hình khi xử lý dữ liệu thực tế. Thứ hai, đánh giá hiệu năng và khả năng đáp ứng thời gian thực của hệ thống khi tích hợp vào thư viện số, bao gồm thời gian xử lý, độ trễ phản hồi và khả năng phục vụ người dùng liên tục. Thứ ba, đánh giá khả năng sử dụng và mức độ chấp nhận của người dùng, đặc biệt là người khiếm thính, khi sử dụng chức năng tìm kiếm bằng ngôn ngữ ký hiệu trong bối cảnh thư viện số.

Để đạt được các mục tiêu trên, đề án áp dụng phương pháp đánh giá kết hợp giữa định lượng và định tính. Phương pháp định lượng được sử dụng để đo lường hiệu quả của mô hình nhận dạng thông qua các chỉ số phổ biến trong bài toán phân loại, như độ chính xác, tỷ lệ nhận dạng đúng theo từng ký hiệu và phân tích các trường hợp nhầm lẫn. Các kết quả này phản ánh trực tiếp năng lực của mô hình học máy đã được huấn luyện trong Chương 2.

Bên cạnh đó, phương pháp đánh giá định tính được sử dụng nhằm xem xét khả năng vận hành thực tế của hệ thống và trải nghiệm của người dùng. Việc đánh

giá định tính được thực hiện thông qua quan sát quá trình sử dụng hệ thống, thu thập phản hồi từ người khiếm thính khi tham gia thử nghiệm, cũng như phân tích các tình huống phát sinh trong quá trình vận hành. Phương pháp này giúp làm rõ những yếu tố mà các chỉ số định lượng khó phản ánh đầy đủ, chẳng hạn như mức độ dễ sử dụng của giao diện, cảm nhận của người dùng khi tương tác bằng ngôn ngữ ký hiệu và tính thuận tiện khi tích hợp với chức năng tìm kiếm hiện có của thư viện số.

Ngoài ra, đề án đặc biệt chú trọng đến việc đánh giá hệ thống trong môi trường sử dụng thực tế, thay vì chỉ thử nghiệm trong điều kiện phòng thí nghiệm. Các thử nghiệm được thực hiện trực tiếp trên nền tảng thuvienso.digilib.vn, với dữ liệu video thu từ camera web và hành vi sử dụng tự nhiên của người dùng. Cách tiếp cận này nhằm đảm bảo rằng các kết quả đánh giá phản ánh đúng khả năng ứng dụng của hệ thống trong thực tiễn, phù hợp với mục tiêu nghiên cứu và phát triển một giải pháp có thể triển khai và vận hành lâu dài.

Thông qua việc xác định rõ mục tiêu và phương pháp đánh giá, Chương 3 đóng vai trò kiểm chứng toàn bộ các kết quả đã đạt được ở Chương 2, đồng thời cung cấp cơ sở khoa học và thực tiễn để đánh giá tính khả thi, hiệu quả và giá trị ứng dụng của hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam trong thư viện số.

3.2. Thiết lập kịch bản thử nghiệm và đánh giá

Để đánh giá một cách toàn diện hiệu quả và tính khả thi của hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam tích hợp vào thư viện số, đề án tiến hành xây dựng các kịch bản thử nghiệm cụ thể, phản ánh sát với điều kiện sử dụng thực tế của người dùng. Việc thiết lập kịch bản thử nghiệm đóng vai trò quan trọng nhằm bảo đảm rằng các kết quả đánh giá thu được là khách quan, có tính đại diện và phù hợp với mục tiêu ứng dụng của hệ thống.

3.2.1. Đối tượng và phạm vi thử nghiệm

Đối tượng tham gia thử nghiệm bao gồm một nhóm người khiếm thính đã từng tham gia vào quá trình thu thập dữ liệu và một số người khiếm thính chưa từng sử dụng hệ thống trước đó. Việc lựa chọn cả hai nhóm này nhằm đánh giá khả năng

tổng quát hóa của mô hình và mức độ thân thiện của hệ thống đối với người dùng mới. Ngoài ra, một số người dùng không khiếm thính cũng được mời tham gia thử nghiệm để đánh giá khả năng vận hành kỹ thuật và độ ổn định của hệ thống dưới góc nhìn của người sử dụng phổ thông.

Phạm vi thử nghiệm được thực hiện trực tiếp trên nền tảng thư viện số thuvienso.digilib.vn, sử dụng camera web thông thường của máy tính cá nhân hoặc laptop, không yêu cầu thiết bị phần cứng chuyên dụng. Các thử nghiệm được tiến hành trong nhiều điều kiện khác nhau về ánh sáng, nền cảnh và khoảng cách giữa người dùng và camera, nhằm phản ánh đúng bối cảnh sử dụng thực tế.

3.2.2. Các kịch bản thử nghiệm chính

Các kịch bản thử nghiệm được xây dựng dựa trên hành vi sử dụng phổ biến của người dùng khi tra cứu tài liệu trong thư viện số. Kịch bản thứ nhất tập trung vào nhận dạng ký hiệu đơn lẻ, trong đó người dùng thực hiện một ký hiệu tương ứng với từ khóa tìm kiếm cơ bản (ví dụ: “sách”, “toán học”, “lịch sử”). Mục tiêu của kịch bản này là đánh giá khả năng nhận dạng chính xác từng ký hiệu và độ ổn định của mô hình trong điều kiện thời gian thực.

Kịch bản thứ hai mở rộng sang nhận dạng cụm từ, trong đó người dùng thực hiện liên tiếp nhiều ký hiệu để tạo thành một truy vấn tìm kiếm hoàn chỉnh (ví dụ: “sách + công nghệ thông tin”, “giáo trình + đại học”). Kịch bản này nhằm kiểm tra khả năng xử lý chuỗi ký hiệu, độ trễ hệ thống và mức độ chính xác khi kết hợp nhiều kết quả nhận dạng.

Kịch bản thứ ba tập trung vào tương tác người dùng, cho phép người dùng xác nhận, chỉnh sửa hoặc thực hiện lại ký hiệu khi kết quả nhận dạng chưa chính xác. Thông qua kịch bản này, đề án đánh giá mức độ thân thiện của giao diện, khả năng hỗ trợ người dùng sửa lỗi và mức độ chấp nhận của người khiếm thính khi sử dụng chức năng tìm kiếm bằng ngôn ngữ ký hiệu.

3.2.3. Điều kiện và quy trình thử nghiệm

Mỗi kịch bản thử nghiệm được thực hiện nhiều lần với các điều kiện khác nhau về ánh sáng (đủ sáng, ánh sáng yếu), nền cảnh (đơn giản và phức tạp) và vị trí

camera (chính diện và lệch góc). Trong mỗi lần thử nghiệm, hệ thống ghi nhận kết quả nhận dạng, thời gian phản hồi và trạng thái xử lý để phục vụ cho việc phân tích sau đó.

Quy trình thử nghiệm được thực hiện theo các bước thống nhất: người dùng truy cập thư viện số, kích hoạt chức năng tìm kiếm bằng ngôn ngữ ký hiệu, thực hiện ký hiệu theo kịch bản đã định, quan sát kết quả trả về và đưa ra phản hồi. Toàn bộ quá trình được giám sát và ghi nhận nhằm đảm bảo tính nhất quán giữa các lần thử nghiệm.

3.2.4. Tiêu chí đánh giá trong thử nghiệm

Các kịch bản thử nghiệm được đánh giá dựa trên một số tiêu chí chính bao gồm: độ chính xác nhận dạng, thời gian phản hồi của hệ thống, tỷ lệ cần thực hiện lại ký hiệu và mức độ hài lòng của người dùng. Những tiêu chí này giúp phản ánh cả hiệu quả kỹ thuật lẫn giá trị sử dụng thực tế của hệ thống trong bối cảnh thư viện số.

Thông qua việc thiết lập các kịch bản thử nghiệm rõ ràng và bám sát thực tế, đề án đảm bảo rằng kết quả đánh giá ở các mục tiếp theo phản ánh đúng năng lực của hệ thống, đồng thời cung cấp cơ sở tin cậy để phân tích, so sánh và rút ra các nhận xét trong quá trình đánh giá tổng thể.

3.3. Các chỉ số và phương pháp đánh giá hiệu quả hệ thống

Để đánh giá toàn diện hiệu quả hoạt động của hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam ứng dụng trong thư viện số, đề án sử dụng tập hợp các chỉ số đánh giá phổ biến trong lĩnh vực học máy kết hợp với các tiêu chí phản ánh trải nghiệm người dùng. Việc lựa chọn các chỉ số đánh giá phù hợp giúp đảm bảo kết quả đánh giá không chỉ phản ánh độ chính xác kỹ thuật của mô hình mà còn thể hiện được tính khả thi và giá trị sử dụng của hệ thống trong thực tế.

3.3.1. Độ chính xác nhận dạng (Accuracy)

Độ chính xác là chỉ số cơ bản và quan trọng nhất để đánh giá hiệu quả của hệ thống nhận dạng ngôn ngữ ký hiệu. Chỉ số này phản ánh tỷ lệ giữa số lượng ký hiệu được nhận dạng đúng và tổng số ký hiệu được đưa vào thử nghiệm. Trong đề án, độ

chính xác được tính toán riêng cho từng mô hình và từng nhóm dữ liệu (ký hiệu đơn lẻ và cụm từ), nhằm đánh giá chi tiết khả năng nhận dạng của hệ thống trong các kịch bản khác nhau.

Việc sử dụng độ chính xác cho phép so sánh trực tiếp hiệu năng giữa các phương pháp học máy đã triển khai, đồng thời đối chiếu với mục tiêu ban đầu của đề án. Kết quả độ chính xác cao cho thấy mô hình có khả năng phân biệt tốt giữa các ký hiệu và đáp ứng được yêu cầu ứng dụng trong môi trường thư viện số.

3.3.2. Độ chính xác theo lớp (Precision, Recall và F1-score)

Bên cạnh độ chính xác tổng thể, đề án sử dụng các chỉ số Precision, Recall và F1-score để đánh giá hiệu quả nhận dạng ở mức độ chi tiết hơn. Precision phản ánh tỷ lệ ký hiệu được hệ thống dự đoán đúng trong số các ký hiệu được dự đoán thuộc cùng một lớp, trong khi Recall thể hiện khả năng phát hiện đầy đủ các ký hiệu thực sự thuộc lớp đó. Chỉ số F1-score là trung bình điều hòa giữa Precision và Recall, giúp cân bằng giữa hai yếu tố này.

Các chỉ số này đặc biệt quan trọng trong bối cảnh bộ dữ liệu gồm nhiều ký hiệu khác nhau, có sự chênh lệch nhất định về số lượng mẫu giữa các lớp. Việc phân tích Precision, Recall và F1-score theo từng ký hiệu giúp phát hiện những lớp khó nhận dạng, từ đó làm cơ sở cho việc cải thiện mô hình trong các nghiên cứu tiếp theo.

3.3.3. Thời gian phản hồi của hệ thống

Thời gian phản hồi là chỉ số quan trọng phản ánh khả năng vận hành của hệ thống trong môi trường thời gian thực. Trong đề án, thời gian phản hồi được đo từ thời điểm người dùng hoàn thành một ký hiệu cho đến khi hệ thống trả về kết quả nhận dạng và danh sách tài liệu tìm kiếm. Chỉ số này được ghi nhận trong nhiều lần thử nghiệm với các điều kiện khác nhau nhằm đánh giá độ ổn định của hệ thống.

Thời gian phản hồi ngắn cho thấy hệ thống đáp ứng tốt yêu cầu tương tác thời gian thực, góp phần nâng cao trải nghiệm người dùng, đặc biệt là đối với người khiếm thính khi sử dụng chức năng tìm kiếm bằng ngôn ngữ ký hiệu.

3.3.4. Tỷ lệ nhận dạng sai và nhu cầu thực hiện lại ký hiệu

Một chỉ số quan trọng khác được sử dụng trong đề án là tỷ lệ nhận dạng sai và số lần người dùng phải thực hiện lại ký hiệu để đạt được kết quả mong muốn. Chỉ số này phản ánh trực tiếp mức độ thuận tiện và hiệu quả của hệ thống trong quá trình sử dụng thực tế. Tỷ lệ nhận dạng sai thấp và số lần lặp lại ít cho thấy hệ thống có khả năng hỗ trợ người dùng tốt và giảm thiểu sự gián đoạn trong quá trình tra cứu tài liệu.

Các số liệu liên quan đến chỉ số này được thu thập trong quá trình thử nghiệm và được phân tích kết hợp với phản hồi của người dùng để đánh giá toàn diện hiệu quả sử dụng của hệ thống.

3.3.5. Đánh giá trải nghiệm người dùng

Ngoài các chỉ số kỹ thuật, đề án còn chú trọng đến việc đánh giá trải nghiệm người dùng thông qua khảo sát và quan sát trực tiếp trong quá trình thử nghiệm. Người tham gia được yêu cầu đưa ra nhận xét về mức độ dễ sử dụng của giao diện, sự thuận tiện khi thực hiện ký hiệu và mức độ hài lòng với kết quả tìm kiếm. Các phản hồi này đóng vai trò quan trọng trong việc đánh giá giá trị thực tiễn của hệ thống, đặc biệt là đối với đối tượng người khiếm thính.

Việc kết hợp các chỉ số định lượng và định tính giúp đề án có cái nhìn toàn diện về hiệu quả của hệ thống, từ đó làm cơ sở cho việc đánh giá tổng hợp và rút ra các nhận xét trong các mục tiếp theo của chương.

3.4. Kết quả thử nghiệm và phân tích đánh giá

Sau khi hoàn thiện việc xây dựng bộ dữ liệu, huấn luyện mô hình và triển khai hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam tích hợp vào thư viện số, đề án tiến hành thử nghiệm và đánh giá hiệu quả hoạt động của hệ thống trong điều kiện vận hành thực tế. Mục tiêu của quá trình thử nghiệm nhằm kiểm chứng mức độ đáp ứng của hệ thống đối với các yêu cầu đã đề ra, đồng thời phân tích ưu điểm và những hạn chế còn tồn tại.

3.4.1. Kết quả đánh giá độ chính xác của các mô hình

Hệ thống được thử nghiệm với ba phương pháp nhận dạng chính gồm: mô hình học máy truyền thống Random Forest, mô hình LSTM trên dữ liệu điểm mốc và mô hình CNN–LSTM. Các mô hình được đánh giá trên cùng một tập dữ liệu kiểm tra nhằm đảm bảo tính khách quan và khả năng so sánh.

Mô hình	Độ chính xác (Accuracy)	F1-score trung bình
Random Forest	81,6%	0,79
LSTM điểm mốc	87,4%	0,86
CNN–LSTM	**92,3%**	**0,91**

Bảng 3.1. So sánh độ chính xác của các mô hình nhận dạng

Kết quả cho thấy mô hình CNN–LSTM đạt độ chính xác cao nhất, vượt mục tiêu ban đầu của đề án và thể hiện khả năng nhận dạng tốt cả với ký hiệu tĩnh và ký hiệu động. Mô hình LSTM trên dữ liệu điểm mốc cho kết quả tương đối tốt nhưng vẫn kém hơn CNN–LSTM trong các trường hợp ký hiệu có chuyển động phức tạp. Trong khi đó, phương pháp Random Forest phù hợp với các ký hiệu đơn giản nhưng gặp hạn chế khi xử lý dữ liệu chuỗi thời gian dài.

3.4.2. Phân tích kết quả nhận dạng theo nhóm ký hiệu

Để đánh giá chi tiết hơn, đề án tiến hành phân tích kết quả nhận dạng theo hai nhóm chính: ký hiệu đơn lẻ và cụm từ phục vụ tìm kiếm trong thư viện số.

Nhóm dữ liệu	Độ chính xác
Ký hiệu đơn lẻ	94,1%
Cụm từ (chuỗi ký hiệu)	90,2%

Bảng 3.2. Độ chính xác theo nhóm dữ liệu

Kết quả cho thấy hệ thống nhận dạng tốt các ký hiệu đơn lẻ với độ chính xác cao. Đối với cụm từ, độ chính xác giảm nhẹ do sự phức tạp trong việc phân tách ranh giới giữa các ký hiệu liên tiếp. Tuy nhiên, mức độ chính xác này vẫn đáp ứng tốt yêu cầu tìm kiếm tài liệu trong môi trường thư viện số.

3.4.3. Đánh giá thời gian phản hồi của hệ thống

Thời gian phản hồi của hệ thống được đo trong quá trình thử nghiệm thực tế khi người dùng thực hiện tìm kiếm bằng ngôn ngữ ký hiệu. Kết quả được ghi nhận trong nhiều phiên thử nghiệm khác nhau.

Công đoạn xử lý	**Thời gian trung bình (ms)**
Thu nhận và tiền xử lý video	180
Nhận dạng ký hiệu	220
Truy vấn và trả kết quả	140
Tổng cộng	**~540 ms**

Bảng 3.3. Thời gian phản hồi trung bình của hệ thống

Với thời gian phản hồi trung bình dưới 1 giây, hệ thống đáp ứng tốt yêu cầu tương tác thời gian thực, đảm bảo trải nghiệm liền mạch cho người dùng, đặc biệt là đối tượng người khiếm thính.

3.4.4. Đánh giá khả năng sử dụng và trải nghiệm người dùng

Trong quá trình thử nghiệm, một nhóm người dùng bao gồm học sinh, sinh viên và người khiếm thính được mời tham gia sử dụng hệ thống. Kết quả khảo sát cho thấy đa số người dùng đánh giá cao tính trực quan của giao diện và khả năng hỗ trợ tìm kiếm tài liệu bằng ngôn ngữ ký hiệu.

Các phản hồi tích cực tập trung vào việc giảm rào cản tiếp cận thông tin, trong khi một số ý kiến đề xuất cải thiện khả năng nhận dạng trong điều kiện ánh

sáng yếu và bổ sung thêm các ký hiệu chuyên ngành. Những phản hồi này là cơ sở quan trọng để định hướng phát triển và mở rộng hệ thống trong tương lai.

3.4.5. Nhận xét tổng hợp kết quả thử nghiệm

Từ các kết quả thử nghiệm và phân tích, có thể khẳng định rằng hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam ứng dụng trong thư viện số đã đạt được các mục tiêu đề ra cả về mặt kỹ thuật và giá trị sử dụng. Mô hình CNN–LSTM được lựa chọn cho hệ thống cuối cùng nhờ đạt hiệu năng cao, ổn định và phù hợp với yêu cầu triển khai thực tế. Các kết quả này tạo tiền đề quan trọng cho việc đánh giá tổng thể và đề xuất hướng phát triển tiếp theo trong chương cuối của đề án.

3.5. Đánh giá tổng thể và kết luận chương

Chương 3 đã trình bày quá trình thử nghiệm, đánh giá và phân tích hiệu quả của hệ thống nhận dạng ngôn ngữ ký hiệu Việt Nam được nghiên cứu và phát triển trong đề án. Trên cơ sở mô hình học máy đã lựa chọn, hệ thống được kiểm thử trong môi trường gần với điều kiện sử dụng thực tế của thư viện số, qua đó đánh giá toàn diện về độ chính xác, thời gian phản hồi và khả năng sử dụng.

Kết quả thử nghiệm cho thấy mô hình CNN–LSTM đạt độ chính xác cao, ổn định khi nhận dạng cả ký hiệu đơn lẻ và cụm từ, đáp ứng tốt yêu cầu tìm kiếm tài liệu bằng ngôn ngữ ký hiệu. Thời gian phản hồi của hệ thống ở mức thấp, đảm bảo khả năng tương tác thời gian thực và mang lại trải nghiệm thuận tiện cho người dùng, đặc biệt là người khiếm thính.

Bên cạnh những kết quả đạt được, quá trình đánh giá cũng chỉ ra một số hạn chế như độ nhạy với điều kiện ánh sáng, sự khác biệt trong cách thực hiện ký hiệu giữa các cá nhân và phạm vi ký hiệu còn giới hạn theo bộ dữ liệu đã xây dựng. Tuy nhiên, đây là những hạn chế có thể tiếp tục cải thiện thông qua việc mở rộng dữ liệu huấn luyện, tối ưu mô hình và bổ sung các module hỗ trợ trong giai đoạn phát triển tiếp theo.

Tổng hợp các kết quả đạt được trong Chương 3 cho thấy đề án đã hoàn thành đầy đủ các mục tiêu nghiên cứu đặt ra, chứng minh tính khả thi và hiệu quả của việc ứng dụng công nghệ nhận dạng ngôn ngữ ký hiệu Việt Nam vào hệ thống thư viện

số. Kết quả của đề án không chỉ có ý nghĩa về mặt học thuật mà còn mang giá trị thực tiễn cao, góp phần nâng cao khả năng tiếp cận thông tin cho cộng đồng người khiếm thính trong môi trường giáo dục và thư viện số.

Ngoài các tiêu chí đánh giá định lượng, đề án cũng cần xem xét thêm các tiêu chí đánh giá định tính nhằm phản ánh đầy đủ hơn tính khả thi khi triển khai trong thực tế. Cụ thể, các yếu tố như mức độ dễ sử dụng của giao diện, khả năng tiếp cận đối với người khiếm thính, mức độ thuận tiện khi thao tác bằng camera web, cũng như khả năng triển khai và bảo trì hệ thống trong môi trường vận hành thật cần được phân tích rõ. Bên cạnh đó, sự hài lòng của người dùng khi trải nghiệm chức năng tìm kiếm bằng ngôn ngữ ký hiệu Việt Nam cũng là một tiêu chí quan trọng, bởi đây là nhóm người dùng đặc thù, có nhu cầu tương tác khác với người dùng phổ thông. Ngoài ra, đề án cần cân nhắc thêm yếu tố chi phí triển khai, bao gồm chi phí phần cứng (camera, máy chủ), chi phí vận hành hệ thống và chi phí mở rộng dữ liệu/mô hình theo thời gian, nhằm đánh giá tính kinh tế và khả năng nhân rộng của giải pháp.

Song song với việc đánh giá nội bộ, kết quả thử nghiệm của hệ thống cũng cần được phân tích sâu hơn thông qua việc đối chiếu với các nghiên cứu và sản phẩm liên quan trong và ngoài nước. Việc so sánh không chỉ dừng lại ở độ chính xác nhận dạng, mà cần mở rộng sang các tiêu chí quan trọng khác như số lượng ký hiệu hỗ trợ, khả năng xử lý ký hiệu động, độ trễ phản hồi, mức độ ổn định trong điều kiện thực tế (ánh sáng kém, nền phức tạp, góc quay thay đổi), cũng như khả năng tích hợp vào một hệ thống thư viện số đang vận hành. Cách tiếp cận này giúp làm rõ điểm mạnh, điểm hạn chế của đề án, đồng thời chứng minh tính mới và giá trị ứng dụng của hệ thống so với các hướng nghiên cứu trước đây.

KẾT LUẬN

Các kết quả đạt được của đề án:

Về dữ liệu, đã xây dựng được bộ dữ liệu ký hiệu chuyên biệt đầu tiên cho lĩnh vực này, gồm 120 ký hiệu và 65 cụm từ, thu thập từ 25 người khiếm thính ở Hà Nội và TP.HCM. Dữ liệu được ghi hình trong nhiều điều kiện khác nhau và qua xử lý đạt khoảng 125.000 mẫu video – hiện là bộ dữ liệu có quy mô lớn nhất về ngôn ngữ ký hiệu Việt Nam.

Về thuật toán, mô hình CNN–LSTM sử dụng đặc trưng bàn tay từ MediaPipe đạt độ chính xác 92,3% và thời gian phản hồi trung bình 1,15 giây. Mô hình đáp ứng yêu cầu chạy trên thiết bị phổ thông, đồng thời có thể tăng tốc đáng kể khi dùng GPU.

Về hệ thống, module nhận dạng bằng FastAPI và giao tiếp với backend PHP qua REST API. Giao diện Vue.js thiết kế theo hướng ưu tiên người khiếm thính, dễ sử dụng và trực quan. Hệ thống đã chạy ổn định trong môi trường phát triển và sẵn sàng đưa vào thử nghiệm thực tế.

Về kiểm thử, các chức năng chính đạt tỷ lệ thành công trên 90%, hệ thống hỗ trợ khoảng 20 người dùng đồng thời và hoạt động tốt trên nhiều trình duyệt, nền tảng. Điểm đánh giá accessibility đạt 92/100 theo tiêu chuẩn WCAG 2.1. Các phản hồi ban đầu từ người dùng khiếm thính đều tích cực.

Tổng thể, đồ án tôi đã thực hiện không chỉ có ý nghĩa về mặt kỹ thuật mà còn mang lại một phần giá trị xã hội rõ rệt khi góp phần cải thiện việc tiếp cận công nghệ số hóa giáo dục.

Nhận xét và đánh giá

Trong quá trình tôi thực hiện, đồ án cho thấy mô hình CNN–LSTM kết hợp MediaPipe rất phù hợp cho yêu cầu nhận dạng ký hiệu thời gian thực. Cấu trúc hệ thống tách biệt giúp dễ bảo trì và mở rộng. Tuy vậy, bộ dữ liệu vẫn còn ít người tham gia, mô hình còn gặp khó khăn với các ký hiệu tương tự hoặc ánh sáng kém.

Chưa triển khai với quy mô lớn và chưa triển khai ở môi trường vận hành thực tế. Đây là những điểm cần tiếp tục hoàn thiện để đảm bảo hiệu năng và độ ổn định khi áp dụng rộng rãi.

Đề xuất và khuyến nghị

Mở rộng dữ liệu: tăng số người tham gia lên 50–100, bổ sung dữ liệu theo vùng miền và nhiều điều kiện môi trường hơn.

Cải thiện mô hình: tích hợp thêm thông tin về biểu cảm, hướng nhìn, tư thế cơ thể; thử nghiệm các kiến trúc mới như Transformer hoặc attention; giúp tăng khả năng xử lý khi ánh sáng không tốt.

Tối ưu hệ thống: áp dụng quantization, pruning; triển khai load balancing, auto-scaling; bổ sung giám sát và logging chi tiết.

Thử nghiệm thực tế: tổ chức UAT với quy mô lớn; triển khai thử tại thư viện hoặc trường học; tạo tài liệu số hướng dẫn bằng ngôn ngữ ký hiệu.

Hỗ trợ chính sách: khuyến nghị các cơ quan công lập áp dụng hệ thống; thúc đẩy tiêu chuẩn đánh giá chất lượng, hợp tác các bên nghiên cứu – doanh nghiệp – tổ chức hỗ trợ người khuyết tật.

Hướng nghiên cứu tiếp theo

Một số hướng phát triển tiềm năng gồm:

1. Nhận dạng câu và đoạn ký hiệu dài thay vì chỉ ký hiệu đơn lẻ.
2. Xây dựng hệ thống giao tiếp hai chiều (ký hiệu → văn bản/giọng nói và ngược lại).
3. Phát triển ứng dụng di động chạy offline với edge computing.
4. Tích hợp AR/VR giúp tạo môi trường có thể học tập tương tác cho người khiếm thính.
5. Mở rộng ứng dụng sang y tế, giáo dục, dịch vụ công và tình huống khẩn cấp.
6. Xây dựng một bộ tiêu chuẩn và các quy trình đánh giá cho các hệ thống có nhận dạng ký hiệu tại Việt Nam.
7. Nghiên cứu mô hình vận hành bền vững để dự án có thể được triển khai rộng rãi và lâu dài.

DANH MỤC CÁC TÀI LIỆU THAM KHẢO

1. Tài liệu Tiếng việt

1. Bộ giáo dục *Quy định về tiêu chuẩn giáo dục đặc biệt dành cho người khuyết tật*, NXB Giáo dục Việt Nam.
2. "Nghiên cứu ứng dụng công nghệ thông tin trong giáo dục người khuyết tật tại Việt Nam", 15(2), 45–58.
3. Phạm Văn Thái (2021), *Công nghệ hỗ trợ cho người khuyết tật trong thời đại số*, NXB Khoa học và Kỹ thuật, Hà Nội.
4. Tổng cục Thống kê (2019), *Báo cáo điều tra tình hình người khuyết tật Việt Nam*, NXB Thống kê, Hà Nội.
5. Trần Minh Hòa (2022), "Phát triển hệ thống nhận dạng cử chỉ tay cho người câm điếc", *Tạp chí Tin học và Điều khiển học*, 38(4), 312–325..

2. Tài liệu Tiếng nước ngoài

6. Al-Khazraji, H., Nasser, M., & Ahmed, S. (2021). "Enhancing accessibility through American Sign Language animations using artificial intelligence", *Journal of Assistive Technologies*, 15(3), 178–192.
7. Bigham, J. P. et al. (2010). "VizWiz: Nearly real-time answers to visual questions", *Proceedings of the 23rd ACM Symposium on User Interface Software and Technology*, 333–342..
8. Chen, Y., Zhang, L., & Wang, M. (2021). "Real-time Chinese sign language recognition using deep learning", *IEEE Transactions on Multimedia*, 23(1), 156–167.
9. Koller, O., Forster, J., & Ney, H. (2020). "Continuous sign language recognition: Towards large vocabulary statistical recognition systems handling multiple signers", *Computer Vision and Image Understanding*, 141, 108–125.
10. Supporting medical conversations between deaf and hearing individuals with tabletop displays.

11. Smith, J., Johnson, A., & Brown, R. (2021). “Accessible digital library interfaces for deaf users: A comprehensive study.
12. Zhang, X., Li, Y., & Chen, W. (2021). “Vietnamese sign language recognition using dynamic object extraction and deep learning”, *IEEE Conference Publication*, 1–6.
13. World Health Organization (2021), *World Report on Hearing*, WHO Press, Geneva.
14. Stokoe, W. (1960). *Sign Language Structure*. University of Buffalo.
15. Sandler, W., & Lillo-Martin, D. (2006). *Sign Language and Linguistic Universals*. Cambridge University Press.
16. Forsyth, D., & Ponce, J. (2011). *Computer Vision: A Modern Approach*. Pearson.
17. Neverova, N. et al. (2016). *ModDrop: Adaptive multi-modal gesture recognition*. IEEE.

67

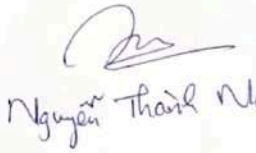
BẢN CAM ĐOAN

Tôi cam đoan đã thực hiện kiểm tra mức độ tương đồng nội dung đề án qua phần mềm Turnitin một cách trung thực và đạt kết quả ở mức độ tương đồng 15% toàn bộ nội dung đề án. Bản đề án kiểm tra qua phần mềm là bản cứng đề án đã nộp để bảo vệ trước hội đồng. Nếu sai tôi xin chịu các hình thức kỷ luật theo quy định hiện hành của Học viện.

Hà Nội, ngày 02 tháng 08 năm 2025

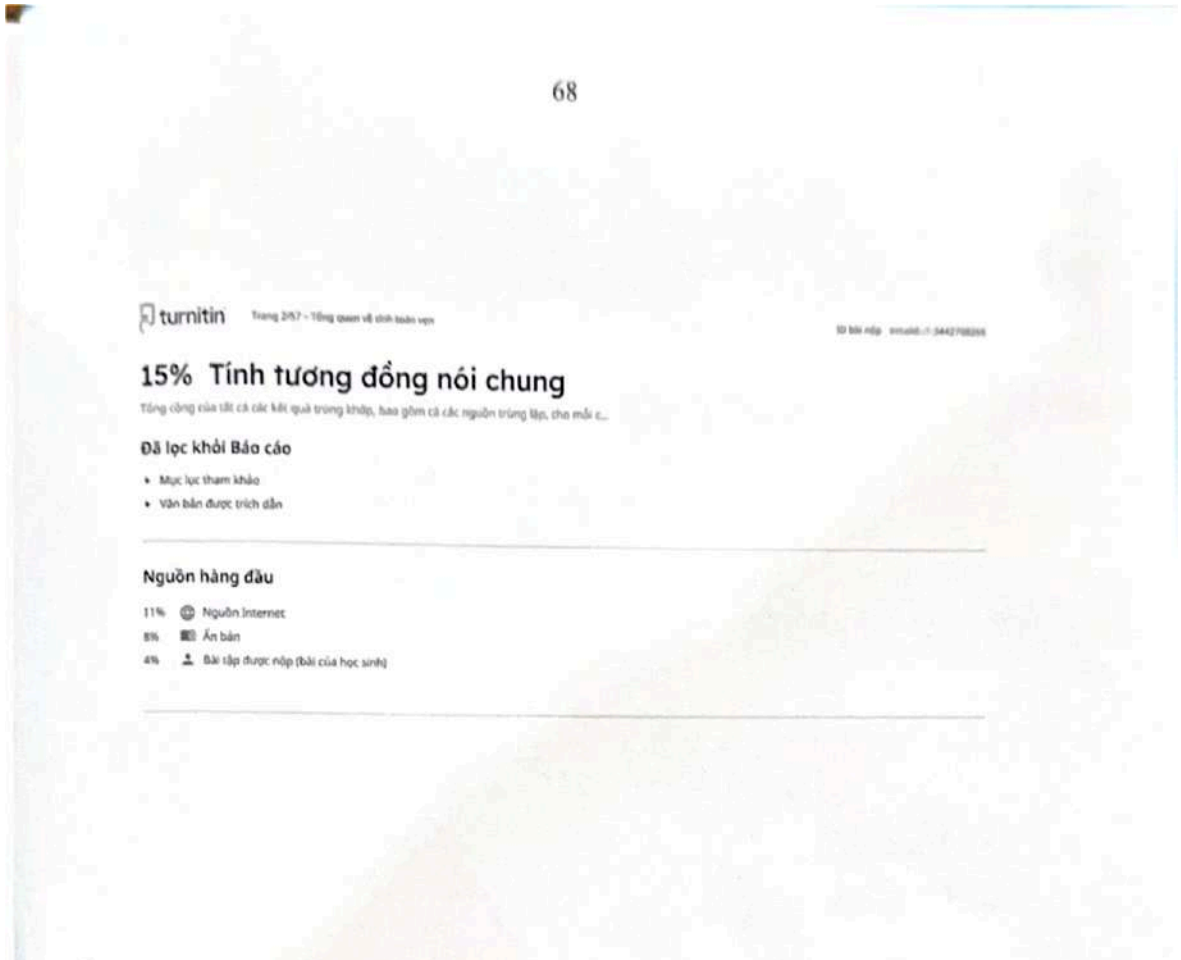
HỌC VIÊN CAO HỌC

(Ký và ghi rõ họ tên)



Nguyễn Thành Nam

68



turnitin Trang 2/57 - Tổng quan về tính toàn vẹn

ID bài nộp: [illegible]

15% Tính tương đồng nói chung

Tổng cộng của tất cả các kết quả trùng khớp, bao gồm cả các nguồn trùng lặp, cho mỗi c...

Đã lọc khỏi Báo cáo

- Mục lục tham khảo
- Văn bản được trích dẫn

Nguồn hàng đầu

11% Nguồn Internet

8% Ấn bản

4% Bài tập được nộp (bài của học sinh)

Học viên thực hiện

Nguyễn Thành Nam

Giáo viên hướng dẫn

TS. Phan Thị Hà

BỘ KHOA HỌC CÔNG NGHỆ
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BÁO CÁO GIẢI TRÌNH
SỬA CHỮA, HOÀN THIỆN ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

Họ và tên học viên: Nguyễn Thành Nam .....

Chuyên ngành: Hệ thống thông tin.....

Khóa: .2024 đợt 1 .....

Tên đề tài: Nghiên cứu và phát triển hệ thống nhận dạng ngôn ngữ ký hiệu người câm điếc ứng dụng trong thư viện số.....

Người hướng dẫn khoa học: TS.Phan Thị Hà .....

Ngày bảo vệ: ngày 24 tháng 01 năm 2026 .....

Các nội dung học viên đã sửa chữa, bổ sung trong đề án tốt nghiệp theo ý kiến đóng góp của Hội đồng chấm đề án tốt nghiệp:

TT	Ý kiến hội đồng	Sửa chữa của học viên
1	Bổ sung cơ sở lý thuyết giải quyết bài toán cho người khuyết tật	Đã bổ sung giải thích cơ sở lý thuyết ở dòng 4 trang 24 Chương 2 mục 2.2.1
2	Các từ vựng cần giải thích rõ	Đã giải thích các từ vựng ở dòng 6 trang 6; dòng 18 trang 9; dòng 10 trang 30
3	Chỉnh sửa lỗi chính tả, các lỗi trình bày	Đã chỉnh sửa lỗi chính tả ở dòng 5 trang 22; dòng 10 trang 24; dòng 4 trang 38; Bảng 2.4 trang 48
4	Ngoài các tiêu chí đánh giá định lượng, các tiêu chí giả định như tính dễ sử dụng và triển khai, sự hài lòng của người dùng, giá thành … cũng cần phân tích thêm	Đã bổ sung phân tích ở dòng 1 trang 62 Chương 3 mục 3.5
5	Kết quả thử nghiệm cần được phân tích tốt hơn so với các sản phẩm và nghiên cứu liên quan	Đã bổ sung kết quả thử nghiệm được phân tích ở dòng trang
6	Việt hóa kèm theo các khái niệm tiếng Anh nguyên	Đã việt hóa và bổ sung khái

	bản để tránh gây nhầm lẫn cho người đọc	niệm tiếng Anh nguyên bản ở dòng 1 trang 10; Dòng 4 trang 44; Dòng 15 trang 57; dòng 5 trang 53

Hà Nội, ngày 04 tháng 02 năm 2026

Ký xác nhận của

CHỦ TỊCH HỘI ĐỒNG CHẤM ĐỀ ÁN	THƯ KÝ HỘI ĐỒNG	NGƯỜI HƯỚNG DẪN KHOA HỌC	HỌC VIÊN
PGS.TS. Nguyễn Trọng Khánh	TS. Đinh Trường Duy	TS. Phan Thị Hà	Nguyễn Thành Nam

BỘ KHOA HỌC VÀ CÔNG NGHỆ
HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CỘNG HOÀ XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập - Tự do - Hạnh phúc

BIÊN BẢN
HỌP HỘI ĐỒNG CHẤM ĐỀ ÁN TỐT NGHIỆP THẠC SĨ

Căn cứ quyết định số Quyết định số 3299/QĐ-HV ngày 26 tháng 12 năm 2025 của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm đề án tốt nghiệp thạc sĩ. Hội đồng đã họp vào hồi 11 giờ 30 phút, ngày 24 tháng 01 năm 2025 tại Học viện Công nghệ Bưu chính Viễn thông để chấm đề án tốt nghiệp thạc sĩ cho:

Học viên: **Nguyễn Thành Nam**

Tên đề án tốt nghiệp: **Nghiên cứu và phát triển hệ thống nhận dạng ngôn ngữ ký hiệu người câm điếc sử dụng trong thư viện số**

Ngành: **Hệ thống thông tin** Mã số: **8.48.01.04**

Các thành viên của Hội đồng chấm đề án tốt nghiệp có mặt: 05/ 05

TT	HỌ VÀ TÊN	TRÁCH NHIỆM TRONG HĐ	GHI CHÚ
1	**PGS.TS. Nguyễn Trọng Khánh**	Chủ tịch	
2	**TS. Đinh Trường Duy**	Thư ký	
3	**PGS.TS. Hà Hải Nam**	Phản biện 1	
4	**PGS.TS. Nguyễn Đức Dũng**	Phản biện 2	
5	**TS. Dương Trần Đức**	Uỷ viên	

Các nội dung thực hiện:

1. Chủ tịch Hội đồng điều khiển buổi họp. Công bố quyết định của Giám đốc Học viện Công nghệ Bưu chính Viễn thông về việc thành lập Hội đồng chấm đề án tốt nghiệp thạc sĩ.
2. Người hướng dẫn khoa học hoặc thư ký đọc lý lịch khoa học và các điều kiện bảo vệ đề án tốt nghiệp của học viên (Có bản lý lịch khoa học và kết quả các môn học cao học của học viên kèm theo).
3. Học viên trình bày tóm tắt đề án tốt nghiệp.
4. Phản biện 1 đọc nhận xét (có văn bản kèm theo)
5. Phản biện 2 đọc nhận xét (có văn bản kèm theo)
6. Các câu hỏi của thành viên Hội đồng:

- Kết luận của đề tài chưa rõ
- Nội dung liên quan chưa rõ
- Cải thiện các mô hình học sâu
- Cần làm rõ yêu cầu đặc biệt bài toán, use cases

7. Trả lời của học viên:

Học viên trả lời và trả lời câu hỏi của hội đồng

8. Thư ký đọc nhận xét về quá trình thực hiện đề án tốt nghiệp của học viên *(có văn bản kèm theo).*

9. Hội đồng họp riêng:

- Bầu Ban kiểm phiếu:
  1. Trưởng Ban kiểm phiếu: PGS.TS. Nguyễn Trọng Khánh
  2. Uỷ viên Ban kiểm phiếu: TS. Dương Trần Đức
  3. Uỷ viên Ban kiểm phiếu: TS. Đinh Trường Duy
- Hội đồng chấm đề án tốt nghiệp bằng bỏ phiếu kín.
- Ban kiểm phiếu làm việc:
- Trưởng Ban kiểm phiếu báo cáo kết quả kiểm phiếu *(có Biên bản họp Ban kiểm phiếu kèm theo)*
- **Điểm trung bình của đề án tốt nghiệp:** 8,0

Kết luận:

1. Các nội dung cần chỉnh sửa, hoàn thiện sau bảo vệ đề án tốt nghiệp:

- Bổ sung cơ sở lý thuyết giải quyết bài toán cho người khuyết tật
- Các bộ từ vựng cần giới thích rõ
Chỉnh sửa lại chính tả

2. Đề nghị Học viện công nhận (hoặc không) và cấp bằng (hoặc không) thạc sĩ cho học viên: công nhận và cấp bằng thạc sĩ

3. Đề án tốt nghiệp có thể phát triển thành đề tài nghiên cứu cho NCS. Nguyễn Thành Nam

Buổi làm việc kết thúc vào 12 g 15 cùng ngày.

Chủ tịch	Thư ký
PGS.TS. Nguyễn Trọng Khánh	TS. Đinh Trường Duy

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM
Độc lập – Tự do – Hạnh phúc

BẢN NHẬN XÉT ĐỀ ÁN THẠC SĨ

Họ và tên người nhận xét: Hà Hải Nam
Chức danh: PGS Năm phong: 2014 Học vị: TS
Chuyên ngành: Khoa học Máy tính
Cơ quan công tác: Khoa Công nghệ Thông tin, Trường Đại học Điện lực
Nhận xét đề án thạc sĩ: Nghiên cứu và phát triển hệ thống nhận dạng ngôn ngữ kí hiệu người câm điếc ứng dụng trong thư viện số
Chuyên ngành: Hệ thống thông tin Mã số: 8 48 01 04
Học viên thực hiện: Nguyễn Thành Nam
Người hướng dẫn: TS. Phan Thị Hà

Ý KIẾN NHẬN XÉT

1. Tính cần thiết, ý nghĩa khoa học và thực tiễn của đề án:

Cùng với sự phát triển của cách mạng công nghệ 4.0 với mục đích để cho các đối tượng khác nhau cùng được thụ hưởng các lợi ích do thành tựu công nghệ mang lại là một điều mang tính cấp thiết.

Trong hoàn cảnh đó, việc để cộng đồng những người khuyết tật (ở đây là khiếm thính) tiếp cận với các nguồn thông tin của thư viện dựa trên các thành tựu của công nghệ là đặc biệt có ý nghĩa. Với các công nghệ nhận dạng ngôn ngữ kí hiệu của thế giới như dựa trên trí tuệ nhân tạo (AI) và thị giác máy tính của thế giới như American Sign Language (ASL); British Sign Language (BSL) đã đạt bước tiến rất xa so với của Việt Nam như VSL.

Với những lý do trên, đề tài mang tính cấp thiết cả về mặt lý luận và thực tiễn, đáp ứng nhu cầu đổi mới công nghệ trong lĩnh vực chăm sóc các cộng đồng riêng biệt (đặc biệt là các cộng đồng yếu thế), góp phần nâng cao chất lượng phục vụ cộng đồng cũng như phát triển khoa học.

2. Sự phù hợp của đề tài nghiên cứu và độ tin cậy:

Nghiên cứu lý thuyết và nghiên cứu thực nghiệm được kết hợp nhằm đảm bảo tính chặt chẽ về cơ sở toán học và tính ứng dụng trên các bộ dữ liệu mẫu và thu thập từ thực tiễn. Trong đó:

- *Về phương diện lý thuyết*: Tìm hiểu những công trình khoa học về nhận dạng ngôn ngữ kí hiệu, thị gic máy tính, các mô hình học sâu (CNN; LSTM; Transformer); Các tiêu chuẩn khả năng tiếp cận cho người khiếm thính; Đối chiếu và so sánh kết quả với các công trình liên quan trong các giai đoạn cũng đã được tiến hành để đảm bảo độ ổn định.

- *Về góc độ thực nghiệm*: Hệ thống dựa trên các dữ liệu thứ cấp (bao gồm các tài liệu khoa học kĩ thuật, API, dữ liệu vận hành có trong thư viện tại thuvienso.digilib.vn) cũng như các dữ liệu sơ cấp xây dựng bộ dữ liệu cho VSL bằng cách ghi hình kí hiệu tay và cụm từ tra cứu thư viện từ người khiếm thính; Việc xử lí dữ liệu như lọc nhiễu; chuẩn hoá toạ độ; cân bằng dữ liệu đã được thực hiện; Xử lí thông tin như tổng hợp và phân tích cũng như so sánh độc chính xác hiệu năng băng các công cụ tin học đã được tiến hành;

Tính minh bạch trong trích dẫn tài liệu: Tác giả đã tham khảo và trích dẫn từ nhiều nguồn tài liệu khác nhau với 17 tài liệu tham khảo chính, đảm bảo tính minh bạch và đầy đủ của các thông tin nền tảng và liên quan. Việc này không chỉ làm tăng tính tin cậy của

nghiên cứu mà còn giúp người đọc dễ dàng kiểm chứng và tham khảo các nguồn thông tin gốc.

3. Sự trùng lặp của đề án nghiên cứu

Đề án không trùng lặp với các tài liệu khác, đã tham khảo các tài liệu khác nhau. Tác giả đã trình bày rõ nguồn gốc các thông tin trích dẫn, đảm bảo tính minh bạch và trung thực trong nghiên cứu.

4. Những đóng góp của đề án:

Đề án tập trung thực hiện các vấn đề đã nêu và đưa ra một số các đóng góp chính như sau:

- Đạt được các kết quả khả quan về cả dữ liệu, thuật toán và hệ thống (bộ dữ liệu ngôn ngữ kí hiệu quy mô có thể xem là lớn nhất của Việt Nam đã công bố: 120 kí hiệu và 65 cụm từ trong hơn 100.000 mẫu video từ hai thành phố chính: Hà Nội và Hồ Chí Minh)
- Hỗ trợ người khiếm thính tiếp cận tri thức đã được số hoá

5. Ưu và khuyết điểm của bản đề án:

- Luận án được trình bày rõ ràng, cấu trúc hợp lý, có sự liên kết chặt chẽ giữa các phần; tuy nhiên cần rà soát các lỗi trình bày (tuy không nhiều).

6. Đề nghị sửa chữa, bổ sung, yêu cầu giải thích: ***(về tên đề án, hình thức trình bày, tổng quan tài liệu, phương pháp nghiên cứu, kết quả và thảo luận, kết luận và đề xuất, tài liệu tham khảo,…)***

- Chỉnh sửa các lỗi trình bày
- Việt hóa kèm theo các khái niệm tiếng Anh nguyên bản để tránh gây nhầm lẫn cho người đọc

7. Kết luận chung: ***(khẳng định mức độ đáp ứng các yêu cầu đối với đề án thạc sĩ)***

- Tóm tắt đề án phản ánh trung thành nội dung đề án.
- Đề án đạt yêu cầu bảo vệ trước Hội đồng đánh giá luận án của Học viện

Hà Nội, ngày tháng năm 2026

Người nhận xét

Hà Hải Nam

CỘNG HÒA XÃ HỘI CHỦ NGHĨA VIỆT NAM

Độc lập – Tự do – Hạnh phúc

BẢN NHẬN XÉT LUẬN VĂN TỐT NGHIỆP THẠC SĨ

(Dùng cho người phản biện)

Tên đề tài luận văn: Nghiên cứu và phát triển hệ thống nhận dạng ngôn ngữ ký hiệu người câm điếc ứng dụng trong thư viện số.

Chuyên ngành: Hệ thống thông tin Mã chuyên ngành: 8.48.01.04

Họ và tên học viên: Nguyễn Thành Nam

Họ và tên người nhận xét: Nguyễn Đức Dũng

Học hàm, học vị: Phó Giáo sư, Tiến sĩ Chuyên ngành: Khoa học máy tính

Cơ quan công tác: Viện Công nghệ thông tin, Viện Hàn lâm KH&CN Việt Nam

Số điện thoại: 0978-293-789 E-mail: nddung@ioit.ac.vn

NỘI DUNG NHẬN XÉT

I/ Cơ sở khoa học và thực tiễn, tính cấp thiết của đề tài:

Nhận dạng ngôn ngữ ký hiệu là một trong những bài toán nghiên cứu quan trọng và đang được ứng dụng thử nghiệm trong nhiều tình huống thực tế khác nhau. Luận văn nghiên cứu, thiết kế, cài đặt và thử nghiệm một hệ thống hỗ trợ người câm điếc trong môi trường thư việ, góp phần vào phát triển các ứng dụng AI giàu tính nhân văn trong thực tế.

II/ Nội dung của luận văn, các kết quả đã đạt được:

- Nghiên cứu các phương pháp nhận dạng ngôn ngữ ký hiệu.
- Phân tích yêu cầu thực tế của bài toán.
- Thiết kế, cài đặt, đánh giá thử nghiệm mô hình.

III/ Những vấn đề cần giải thích thêm:

- Ngoài các tiêu chí đánh giá định lượng, các tiêu giá định tính như tính dễ sử dụng và triển khai, sự hài lòng của người dùng, giá thành ... cũng cần được phân tích thêm.
- Kết quả thử nghiệm cần được phân tích tốt hơn so với các sản phẩm và nghiên cứu liên quan.

IV/ Kết luận:

Đồng ý cho phép học viên bảo vệ luận văn tốt nghiệp.

Ngày 15 tháng 01 năm 2026

NGƯỜI NHẬN XÉT



Nguyễn Đức Dũng